Using the Alternative Minimum Tax to Estimate the Elasticity of Taxable Income for High Earners*

Ali Abbas†

This draft: November 5th, 2020
Latest draft: Click here

Abstract

High-income taxpayers contribute a large proportion of the personal income tax revenue in the United States. Their taxable income response to the rate of taxation, measured by the elasticity of taxable income with respect to the net-of-tax rate is a core parameter for estimating the efficiency costs of taxation and optimal top marginal tax rates. Studies that use bunching methods to study taxable income response at kinks in the income tax schedule for estimating this parameter have shown no evidence of high-income responses around the top kink in the regular, federal income tax schedule. I argue that this schedule does not identify the actual, tax-related incentives that apply to high-income individuals. At the federal level, high earners are subject to a combination of the regular income tax and the Alternative Minimum Tax. I use annual income tax codes and publicly available samples of Internal Revenue Service individual income tax return data from 1993-2011 to characterize the combined schedule for each taxpayer. I discover previously undetected bunching at the top kink in this schedule and use it to estimate the average elasticity for high earners to be between 0.15 to 0.28. This estimate implies a lower bound on efficiency cost of 22 cents to 45 cents per dollar of additional tax revenue collected, signaling an upper bound on optimal marginal tax rates of 70 percent to 82 percent. My approach makes a unique methodological contribution to the literature: I show that the location of the top kink for each taxpayer on the combined schedule varies across the distribution of taxable income. This generates novel variation in marginal tax rates that is separable from variation in taxable income, allowing me to mitigate a key endogeneity concern associated with the use of bunching estimators on fixed kink points.

Keywords: Elasticity of taxable income, personal income tax, alternative minimum tax, bunching
JEL Classification: H21, H24, H26

---

* I am grateful to Michael Lovenheim, Nancy Chau, Ravi Kanbur, Ariel Ortiz-Bobea, Zhuan Pei, and Douglas Miller for their comments. This research has also benefited from comments received at seminars at the Economics Department and the School of Applied Economics and Management at Cornell University. I would like to give a special thanks to participants at the National Tax Association Annual Conference 2019 and the International Institute of Public Finance Annual Congress 2020 for their valuable feedback.

† Charles H. Dyson School of Applied Economics & Management, Cornell University. Email: aa2378@cornell.edu

I. Introduction

Personal income taxation is a key source of revenue for financing public goods and redistributive schemes. However, non-lump sum personal income taxes alter the after-tax price of labor, incentivizing individuals to change their labor supply (Pencavel, 1986; Hausman, 1991; MaCurdy et al., 1991), shelter earned income from taxation by consuming more tax-deductible items such as healthcare and housing (Glaeser and Shapiro, 2003), or illegally under-declare income (IRS, 2016). These responses can generate deadweight loss in the economy and cause tax revenue loss if the size of the taxable economy shrinks, making the marginal tax rate structure a highly debated policy and political issue. These debates rest in large part on how high-income individuals respond to changes in the marginal tax rate in the top income bracket.

I focus on high income taxpayers because of three reasons. First, the top quintile (percentile) of income earners by households in the United States contribute approximately 88 percent (38 percent) of personal income tax revenue (Tax Policy Center, 2019), so taxable income responses in this group can have substantial revenue consequences.

Second, the response to tax changes in the right-tail of the income distribution can itself be higher relative to the rest of the distribution, given high earners' access to diverse financial strategies including income-shifting across tax bases, retiming of income realization, and the increased use of itemized deductions such as home mortgage and business expense allowances (Saez, Slemrod and Giertz, 2012). For example, part of the income of high earners such as executives could be in the form of stock options, which face lower marginal tax rates on the capital gains schedule, as compared to the top marginal tax rate on the income tax schedule (Hanlon et al., 2005). Taxpayers can also retime capital gains realizations, as documented Goolsbee (2000b), Parcell (1995) and Samartino and Weiner (1997). Increased bargaining power of these taxpayers such as top executives can also allow them to substitute taxable income with non-taxable fringe benefits at work, such as improved work facilities and better healthcare benefits (Piketty et al., 2014). Top earners also have access to sophisticated tax planning services, and self-employment income that is not reported by third parties creates space for tax evasion (Slemrod, 2007; Hurst et al., 2010).

Third, the magnitude of the ETI parameter for high earners is highly contested in the public finance literature. However, it is the hypothesized high responsiveness of top earners and the sensitivity of revenues to the high-income tax base that served as a factor in the Reagan tax cuts of 1981 that reduced the top marginal tax rate from 70 percent to 50 percent; and again, in 1986 when the top rate was decreased to 28 percent. Contested views on the responsiveness of the high-income tax base continue to pervade the policy and political discourse.

Measuring taxpayer responses along the labor, avoidance, and evasion margins separately is infeasible due to the inability to observe all of the dimensions of behavior. Instead, Feldstein's (1999) canonical model shows that all such margins of taxpayer responses that affect taxable income and generate deadweight burden are captured by the elasticity of taxable income (ETI) with respect to the net-of-tax rate.[1] This makes the ETI a sufficient statistic for estimating efficiency costs of income taxation and conducting welfare analyses[2], assuming no transfer costs[3] of sheltering (Chetty, 2009) and no fiscal externalities[4] (Slemrod, 1998; Saez, 2004). This makes the ETI a core parameter in the public economics literature. Previous work on estimating the magnitude of high-income taxpayer responsiveness has generated mixed results. For example, Feldstein (1995) estimates the ETI for high earners to be as high as 1.7, while others studying bunching behavior around the top kink in the regular income tax schedule have found no response (Saez, 2010; Mortenson & Whitten, 2016).

In this paper, I employ a bunching estimator to study the responsiveness of top earners in the United States to changes in marginal tax rates by using the intersection of the Alternative Minimum Tax (AMT) and regular income tax schedules. To date, no prior research has studied the combined schedule for estimating the ETI for high earners, or otherwise.

Existing literature on estimating the ETI for high earners in the United States predominantly uses two approaches. The first approach uses taxable income responses related to tax reforms that change top marginal tax rates to estimate the ETI. However, rising inequality that differentially affects secular growth rates in different parts of the taxable income distribution presents a challenge, since it becomes difficult to disentangle the effect of secular income growth on taxable income from the effect of tax rates. Estimates vary significantly, in the range of 0 to 1.7. Initial estimates tended to be high (Lindsey, 1986; Feldstein, 1995). More recent studies that have attempted to isolate variation in taxable income from secular income growth have generated lower estimates. These issues and the relevant literature are discussed in more detail in Section II.

[1] The net-of-tax rate is the post-tax, take-home portion of the marginal dollar earned by a taxpayer.
[2] This is because in the canonical model, the marginal private value of sheltering an additional dollar of income and the marginal social value of earning an additional dollar of income are both pegged to the tax rate.
[3] Chetty (2009) shows that transfers costs to taxpayers of avoiding or evading taxes can be offset by a positive externality on other agents. For example, penalties paid to the government due to tax evasion are redistributed; and an increase in deductible charitable contributions generates positive externalities for other agents in the economy.
[4] Tax revenue losses due to income shifted from one stream can be partially offset by taxation in another stream.

The second approach uses bunching methods on cross-sectional data to avoid identification issues created by secular income growth. The use of this approach in the United States involves estimating the magnitude of bunching at kinks in the regular, federal income tax schedule. This bunching is presumably a result of taxpayers strategically locating on the side of the kink that offers the lower marginal tax rate. This bunching is compared to the tax rate differential around the kink point to estimate the ETI. This approach has revealed no high-income responses around the top kink in the regular, federal income tax schedule (Saez, 2010; Mortenson and Whitten, 2016). On the other hand, the estimated ETI for low-income individuals in these studies is higher, in the range of 0.1 - 0.3, raising the question of why bunching estimators have failed to show evidence of economically significant elasticities for high earners who have more margins along which they can respond.

I argue that the federal, regular income tax schedule used by previous bunching studies does not identify the actual, tax-related incentives that apply to high-income individuals. At the federal level, high earners respond to a combination of the regular income tax and the AMT. The AMT is a concomitant income tax schedule with its own definition of taxable income and marginal tax rates. The purpose of the AMT is to ensure that high-income taxpayers do not take disproportionate advantage of deductions – which reduce taxable income – offered by the regular income tax schedule. It disallows major deductions such as personal exemptions, the standard deduction, and important itemized deductions such as the state and local tax (SALT) deduction, and miscellaneous deductions used primarily by business owners.[5] By redefining taxable income, the AMT causes a part of earned income to be counted as taxable that is otherwise sheltered from taxation on the regular income tax schedule. However, the schedule provides a substantial fixed deduction that prevents low- to middle-income taxpayers from being affected by it.

Taxpayers separately calculate their income tax liabilities on the regular income tax and the AMT schedules and are liable for the higher of the two taxes. The effective schedule is, therefore, the upper envelope of the interaction of the two schedules. The intersection kink – where the two schedules cross – is the top kink in the combined schedule. I find that between 1993-2011, less than 0.5 percent of taxpayers with real 2017 adjusted gross income (AGI) of less than $100,000 were subject to the combined AMT-regular schedule. For real AGIs between $100,000 to $200,000, this rate rises to approximately 3 percent. Amongst taxpayers with real AGI above $300,000, more than 65 percent were subject to the combined schedule, implying a large

[5] The AMT also partially disallows medical and dental deductions, accelerated depreciation, and deductions on home mortgage interest on non-primary property, among others.

proportion of high earners for whom the correct schedule to analyze is the combined, rather than just the regular income tax schedule.

Studying taxpayer behavior around the top kink in the regular income tax schedule in isolation can reveal low ETI estimates for two reasons. First, the top kink in the combined schedule does not systematically align with the top kink in the regular schedule. Studying bunching behavior only around the latter will introduce measurement error and bias estimates of the ETI downward, because the top kink in the regular schedule does not affect taxpayers who are subject to the combined schedule. For these taxpayers, strategic decision-making occurs around the top kink in the combined schedule. Second, the difference in marginal tax rates on the two sides of the top kink in the regular schedule can be too small to elicit a substantial bunching response, even in the absence of the combined schedule. Larger tax rate differentials around kinks create stronger incentives for taxpayers to bunch on the side of the kink point that offers a lower tax rate (Chetty et al. 2011). Such differentials exist on the combined schedule. The marginal tax rate changes from 28 percent to the left of the top kink on the combined schedule to approximately 38 percent to its right, as compared to the approximately 36 to 39 percent (33 to 35 percent) change across the top kink in the regular income tax schedule between 1993-2002 (2003-2011).

Using annual federal income tax codes and publicly available Internal Revenue Service (IRS) income tax data from 1993-2011, I construct the two piecewise linear functions associated with the regular income tax and the AMT schedules for each taxpayer in each tax year, adjusting for taxpayer-level deductions. Across years, the shape of the two tax functions is determined by legislative rules related to the size of income tax brackets and corresponding marginal tax rates. Within years and across individuals, the location of the intersection kink is determined by the amount of deductions allowed by the regular income tax, relative to the AMT. Once constructed, I solve these tax functions for each taxpayer in my sample to find the complete set of intersection kinks. Since the location of the intersection kinks varies for each taxpayer, I recenter these kink points and overlay the observed distribution of taxable income to provide visual evidence of the aggregate bunching response of high earners around the top kink. I estimate this excess mass as compared to an estimated counterfactual density[6] that is a fitted polynomial of the seventh order and use it in a standard bunching estimator to measure the ETI for high earners.

I also test the robustness of my estimates by using weaker assumptions for the functional form of the counterfactual density. Earlier bunching studies that have estimated the ETI for high

[6] The counterfactual density is the underlying distribution if there was no kink point and therefore, no differential taxpayer response to changing tax rates.

earners have made strong functional-form assumptions. For example, the counterfactual density is assumed to be linear in Saez (2010) and a polynomial of order seven in Chetty et al. (2011). I use the method proposed in Bertanha et al. (2020) to estimate non-parametric bounds on the ETI, by using the area of the observed distribution as a constraint on the counterfactual density to restrict its range of slopes in the bunching region.

I also use this setting to counter a methodological weakness of standard bunching estimators. The location of the top kink in the combined schedule varies across the distribution of taxable income for each taxpayer, providing novel variation in marginal tax rates that is separable from variation in taxable income. This feature allows me to mitigate an important endogeneity concern associated with the use of bunching estimators on fixed kink points (Blomquist and Newey, 2017; Bertanha et al., 2020).

The estimated average ETI for high earners at the top kink of the combined schedule is 0.15. This estimate is bounded below at 0.12 and above at 0.17. The estimated ETI for high earners rises to 0.20 for taxpayers who are unaffected by the additional complexity of the capital gains schedule. The Jobs and Growth Tax Relief Reconciliation Act (JGTRRA) of 2003 was followed by annual increases in the AMT fixed deduction amount, pushing the intersection kink to higher income levels where taxpayers are plausibly more responsive to changes in marginal tax rates. For taxpayers unaffected by the additional complexity of the capital gains schedule between 2003-2011, the estimated ETI is 0.28. High earners' responsiveness to marginal tax rates increases over time, with taxpayers earning any self-employment income responding more than others. I apply simplified formulas in the literature that use the estimated ETI parameter in conjunction with marginal tax rates and the shape of the income distribution to estimate efficiency costs and optimal top marginal tax rates, as discussed in Section VII. Intuitively, higher taxpayer responsiveness generates larger distortions in the economy leading to higher efficiency costs and lower optimal top marginal tax rates. My estimates for the average ETI for high earners of 0.15 - 0.28 imply efficiency costs ranging from 22 cents to 45 cents per dollar of additional tax revenue collected. Estimated optimal top marginal tax rates lie between 82 percent and 70 percent. In the presence of transfer costs and fiscal externalities, these estimates serve as upper bounds on estimated efficiency costs and lower bounds on optimal top marginal tax rates (Chetty, 2009).

I make three key contributions to the literature. First, I account for the interaction of the regular income tax and AMT schedules in the United States to show evidence of substantial bunching around the top kink in the combined schedule, resulting in elasticities of 0.15 to 0.28. In contrast, earlier studies showed no response at the top kink of the regular income tax schedule (Saez, 2010; Mortenson and Whitten, 2016). To date, no prior study has studied the combined schedule for estimating the ETI for high earners, or otherwise.

The second contribution that I make is methodological. I provide a unique setting that mitigates recent endogeneity concerns related to the use of bunching methods on kink points fixed in taxable income (Blomquist and Newey, 2017; Bertanha et al., 2016, 2020). Since fixed kinks at which marginal tax rates change are jointly determined with taxable income, observed taxable income is likely correlated with unobserved heterogeneity. Intuitively, it is plausible that individuals select into particular bins of the income distribution not as a result of strategic responses to marginal tax rates but because of some underlying preferences for those income levels. If this occurs, then observed bunching (or troughs) in the taxable income distribution might reflect preferences rather than strategic decision-making related to tax rates, causing bias in the estimation of the ETI of unknown direction. However, in the setting that I leverage, the top kink in the combined schedule varies for each taxpayer across taxable income generating a distribution of top kinks, as illustrated in Section III. This is unique feature of the combined schedule weakens the correlation between taxable income and unobserved heterogeneity, increasing confidence in the ability of my estimator to estimate an unbiased ETI parameter.

Third, I contribute to the small literature on the AMT by providing the only estimates on taxpayers' responses to the AMT. Previous literature in this area has specifically focused on forecasting the coverage and revenue impact of evolving AMT laws of the early 2000s (Burman et al., 2003), its impact on average marginal tax rates (Feenberg and Poterba, 2003), and the role of the AMT as a fiscal stabilizer (Galle and Klick, 2011). However, the AMT has not been leveraged to assess taxpayer behavior and its impact on efficiency and tax revenue.

From a policy perspective, my results point to optimal top marginal tax rates that are higher than prevailing top marginal tax rates. Higher ETI for self-employed individuals confirms the previously documented positive relationship between the absence of third-party reporting and higher tax avoidance behavior. And a comparison of the relationship between bunching responses and the size of the marginal tax rate change around kinks suggests that a larger number of income tax brackets with smaller marginal tax rate changes across brackets will reduce taxable income responses, leading to lower efficiency costs of taxation.

II. Prior Literature

I contribute to the literature on estimating the elasticity of taxable income (ETI) with respect to the net-of-tax rate for high earners. The ETI measures the taxable income response of taxpayers to changes in marginal tax rates. As discussed in Section I, this parameter can be a sufficient statistic for estimating efficiency costs and optimal top marginal tax rates (Feldstein, 1999) under no transfer costs (Chetty, 2009) and no fiscal externalities (Slemrod, 1998; Saez, 2004). In particular, the large share of tax revenue generated by high-income taxpayers and their greater hypothesized ability to respond to changes in tax rates makes studying the ETI of high earners extremely important.

The core challenge with estimating the ETI is related to the endogeneity of tax rates, since taxable income and tax rates are jointly determined. As taxable income rises, the marginal tax rate that the taxable income is subject to increases under a nonlinear schedule. This makes it difficult to disentangle variation in marginal tax rates from variation in taxable income. Prior literature has predominantly used two methods to address this endogeneity concern. The first approach to estimating the ETI for high earners leverages tax reforms that introduce plausibly exogeneous changes in marginal tax rates. The major, federal tax reforms that have been studied in the literature include the Economic Recovery Tax Act (ERTA) of 1981, the Tax Reform Act (TRA) of 1986, the Omnibus Reconciliation Acts (OBRA) of 1990 and 1993 and the American Taxpayer Relief Act (ATRA) of 2012.[7] I compare the estimates from some of the seminal studies using tax reforms in Figure 1.[8] Panel A sorts these estimates by publication year of the study, and shows that studies using tax reforms have found a wide range of estimates, ranging from 0 to 1.7, with more recent studies finding lower estimates of the ETI for high-income taxpayers. Panel B sorts the studies by the median year of analysis considered in each study.[9] Details on the studies represented in Figure 1 are provided in Table A in the Appendix in the Appendix.

Initial estimates using this approach tend to be high (Lindsey, 1986; Feldstein, 1995). Lindsey (1986) and Feldstein's (1995) identification strategies rely on secular growth rates of real income being the same for the groups being compared. If these growth rates vary across groups due to non-tax related reasons, then taxable incomes of taxpayers in high-income groups would be

[7] ERTA 1981 and TRA 1986 reduced the top marginal tax rate from 70 percent to 50 percent, and from 50 percent to 28 percent, respectively. OBRA 1990 and 1993 increased the top marginal tax rate from 28 percent to 31 percent, from 31 percent to 39.6 percent. ARTA 2012 increased the tax rate from 35 percent to 39 percent.

[8] If a study has multiple estimates for the ETI of high earners, I average the estimated ETI.

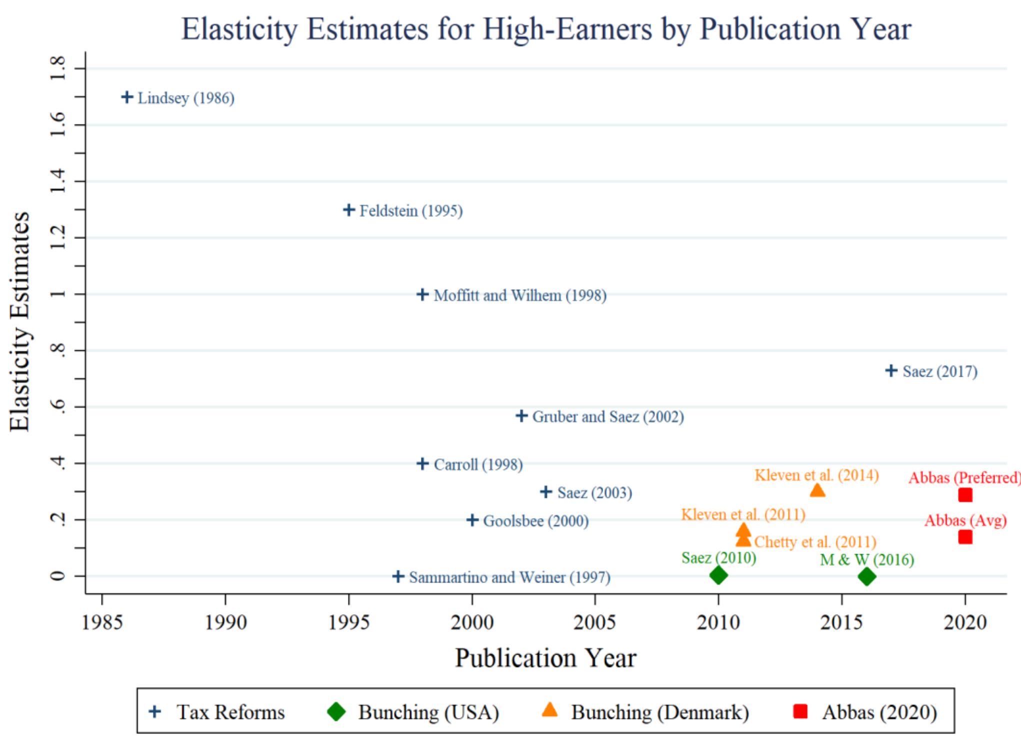
[9] It shows how the ETI estimates related to tax reforms in the 1980s were higher than those that were introduced later. It is possible that structural ETI was higher in the 1980s, due to features of tax audit system, or the specific aspects of tax reforms in this time period.

different from taxpayers in low-income groups across time, even in the absence of tax changes. This differential income growth is well documented. Saez and Zucman (2020) find that between 1980-2018, the national income share of the top one percent grew by 2 percent per year, compared to an annual, average growth rate of 0.2 percent for the bottom 50 percent of the income distribution. The higher, secular growth in the income share of high-income groups would bias the estimate of ETI for high earners in Lindsey (1987) and Feldstein (1995) upwards, plausibly leading to the high estimates observed in these studies. To deal with this issue, most tax reform studies conducted after 1995 controlled for time trends and exploited instrumental variables to disentangle variation in tax rates from variation in taxable income, producing smaller estimates (Gruber and Saez, 2002; Saez, 2003).[10] As shown in Panel A of Figure 1, the estimated ETI of high-income taxpayers is lower, in the range of 0 to 1 from year 1997 onwards.

The second approach attempts to avoid these identification issues by employing bunching methods to estimate the ETI using cross-sectional income tax data. This approach involves overlaying the observed taxable income distribution across a stable, income tax schedule. Observed bunching in this distribution around kinks in the tax schedule plausibly reflects strategic responses of taxpayers in terms of taxable income, with taxpayers locating on the side of the kink where the marginal tax rate is lower. The excess mass in the distribution captures this strategic response and is compared to the change in marginal tax rates at the kink to estimate the ETI. While studies using tax reforms have found a wide range of estimates, bunching methods have found no taxable income response at the top kinks of the income tax schedule in the United States (Saez 2010; Mortenson and Whitten, 2016). In Panel A of Figure 1, I compare estimates for the ETI of high earners in the US with estimates from bunching studies conducted using Danish tax data. It is notable that bunching estimates are non-zero for Danish data, in the range of 0.1 to 0.3. In fact, recent estimates for the ETI of high earners in China, not included in the figure, stand at 0.41 (He et al., 2018).

[10] For a detailed discussion of other identification issues related to studies using tax reforms, review Saez, Slemrod and Giertz (2012).

Panel A



Panel B

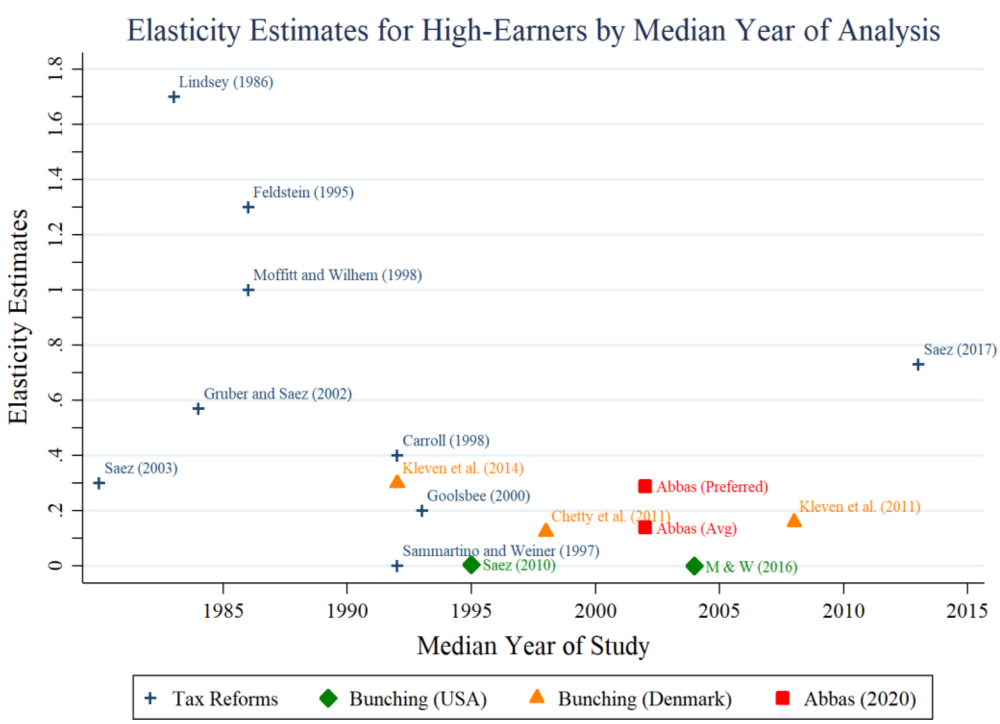


Figure 1: Historical Estimates of the ETI for High-Income Taxpayers

Notes: The scatterplot in Panel A identifies high-earner ETI estimated in previous literature, sorted by Publication year. Panel B contains the same studies sorted by the median year of the analysis sample used by each study. Studies are divided into four types: non-bunching that use other methods, bunching studies in the USA, bunching studies in Denmark, and estimates obtained in this study.

There are two potential reasons for the difference between estimates of the observed ETI for high earners in the United States and in other countries. First, it is possible that high-income taxpayers in the United States simply do not respond to top kinks in the income tax schedule, as compared to their global counterparts, due to reasons including a lack of salience of the top kinks and low structural elasticities. It is also plausible that earlier studies in the US do not consider relevant features of the income tax code when measuring taxpayer bunching responses, resulting in omitted variable bias that introduces downward bias in these estimates. In this paper, I argue that the federal, regular income tax schedule used by previous bunching studies is insufficient to map the effective tax schedule that applies to high-income individuals. At the federal level, high earners are more likely to be subject to both the regular income tax and the Alternative Minimum Tax (AMT) schedules. The AMT has its own marginal tax rates and allowable deductions. Taxpayers separately calculate their income tax liabilities on the regular income tax and the AMT schedules and are liable for the higher of the two taxes. The effective schedule for taxpayers, therefore, is the upper envelope of the interaction of the two schedules. I discuss the structure of the combined schedule in more detail in Section III.

By considering taxpayer behavior along the combined schedule, I provide evidence of previously undetected bunching at the top kink of the combined schedule, in contrast to studies using the top kink in the regular income tax schedule, resulting in higher estimates of the ETI for high-income taxpayers, as shown in Figure 1. My estimates of 0.15 to 0.28 are more in line with bunching studies conducted in other countries. To the best of my knowledge, this is the first paper to study taxpayer responses to the combined schedule, specifically in relation to the AMT. Earlier literature has forecasted the coverage and revenue impact of evolving AMT laws of the early 2000s (Burman et al., 2003), assessed the AMT's impact on average marginal tax rates (Feenberg and Poterba, 2003), and studied the role of the AMT as a fiscal stabilizer (Galle and Klick, 2011). However, the AMT has not been leveraged to assess taxpayer behavior and its impact on efficiency costs of taxation.

I also provide a unique setting that mitigates recent endogeneity concerns related to the use of bunching methods on kink points fixed in taxable income. By providing a unique setting where the location of the top kink in the personal income tax schedule varies across taxpayers, I am able to disentangle variation in marginal tax rates from variation in taxable income to better address endogeneity concerns related to traditional bunching estimators. Earlier bunching studies use tax schedule kinks that are fixed in taxable income in a given tax year. For a single budget set, variation in tax rates across the budget set occurs with variation in taxable income as well as with variation in preferences. The correlation of taxable income and underlying preferences makes it challenging to distinguish the taxable income elasticity from unobserved heterogeneity (Blomquist and Newey, 2017). Intuitively, it is impossible to know if an individual chooses to locate at a kink because of tax rate variation or due to underlying preferences. The variation in the location of the top kink in the combined schedule across high-income taxpayers generates multiple budget sets, limiting exposure to such selection bias by delinking variation in marginal tax rates from variation in taxable income.

The next section provides a detailed overview of the AMT, focusing on its features that interact with the regular income tax schedule to give rise to the combined, effective personal income tax schedule at the federal level in the United States.

III. Conceptual Framework for the Combined Schedule

In this section, I assess the coverage of the federal, combined AMT-regular tax schedule and unpack specific features of the regular income tax and AMT schedules that give rise to the top, intersection kink in the combined schedule. I show how the location of the top kink in the combined schedule varies across taxpayers, potentially creating a downward bias in earlier estimates of the ETI for high-income taxpayers that only looked at the top kink in the regular schedule. I also discuss how the variation in the location of the top kink in the combined schedule can be used to address endogeneity concerns associated with the earlier use of bunching estimators on fixed kink points, and provide evidence for this variation.

The AMT reduces the ability of high-income taxpayers to shelter income from taxation with the use of deductions. The AMT and the regular income tax systems function in parallel to each other. Taxpayers calculate their income tax liability using both the regular income tax form (Form 1040), as well as the AMT form (Form 6251). Once taxpayers have calculated personal income tax liabilities based on both schedules, they must pay the higher of the two amounts, represented by the upper envelope of the interaction of the two schedules.

The number of taxpayers who are subject to the upper envelope of the combined AMT-regular tax schedule increases at higher income levels. For example, I find that approximately 0.03 percent of the population of taxpayers with real adjusted gross income (AGI) less than $50,000 faces the combined AMT-regular schedule. On the other hand, 47 percent (60 percent) of taxpayers with real AGI greater than $200,000 ($300,000) face the AMT. Table 1 provides the fraction of taxpayers who are subject to the combined schedule by real 2007 AGI brackets. The table also provides this breakdown for the subpopulation that submitted Form 6251, the form used to report AMT liability. Since taxpayers submitting this form already expect to be subject to the combined schedule, the fraction of taxpayers who are subject to the AMT, conditional on submitting Form 6251 is close to 100 percent at high income levels.

Table 1: Fraction of Taxpayers Facing the Combined AMT-Regular Tax Schedule

Real (2007) AGI Brackets in '000s of $	**% of taxpayers facing combined schedule**	**% of taxpayers facing combined schedule, conditional on submitting Form 6251**
less than 50	0.03	1.64
50 to 100	0.35	8.36
100 to 200	3.11	23.56
200 to 300	33.09	62.63
300 to 400	63.05	94.97
400 to 500	68.11	97.58
more than 500	57.2	93.5

Studying taxable income response for high earners around the top kink in the regular income tax schedule without accounting for the AMT and the presence of the combined tax schedule can affect the estimation of their ETI with the use of bunching methods through two channels. First, the top kink in the combined schedule does not systematically align with the top kink in the regular income tax schedule. Studying bunching behavior only around the latter will introduce measurement error and bias estimates of the ETI downward, because the top kink in the regular schedule does not affect taxpayers who are subject to the combined schedule. For these taxpayers, strategic decision-making occurs around the top kink in the combined schedule.

Second, taxpayers on the margin are incentivized to locate on the side of the kink offering the lower marginal tax rate. In fact, as shown by Chetty et al. (2011), the utility loss at higher marginal tax rates can justify higher adjustment costs for taxpayers to relocate with respect to the kink point. The top kink on the combined schedule provides a more substantial jump in marginal tax rates relative to the regular schedule at high income levels. Specifically, the marginal tax rate at the top kink in the combined schedule increases from 28 percent to approximately 39 percent (35 percent) between 1993-2002 (2003-2011). Compare this to changes at the top kink in the regular schedule, where the marginal tax rate increases from 36 percent to approximately 39 percent between 1993-2002, and from 33 percent to 35 percent between 2002-2011. The intersection kink, therefore, becomes a valuable device for assessing taxpayer responsiveness to changing marginal tax rates. Below, I discuss the features of the regular schedule, the AMT schedule, and their interaction that results in misalignment of kinks between the regular and combined schedules, and the larger changes in marginal tax rates at the top kink of the combined schedule relative to the regular schedule.

The AMT differs from the federal, regular income tax schedule in three distinct ways, related to taxable income brackets, marginal tax rates, and the definition of taxable income.[11] First, the regular income tax and AMT schedules contain taxable income brackets of different sizes. The regular income tax schedule had five brackets between 1993-2001, and then six brackets between 2002-2012. In contrast, the AMT schedule contains two statutory taxable income brackets. However, a fixed deduction provided by the AMT is phased out at high income levels, causing the AMT to have four distinct effective taxable income brackets. Second, both schedules exhibit different marginal tax rates corresponding to each taxable income bracket. As an example, the differences in taxable income brackets and corresponding marginal tax rates for married joint filers in tax year 2000 are provided in Table 2.

Table 2: Taxable Income Brackets and Marginal Tax Rates for the Regular Tax and AMT in Year 2000

Regular Taxable Income (MFJ)	**Tax Rates**
$0 - $43,850	15%
$43,850 - $105,950	28%
$105,950 - $161,450	31%
$161,450 - $288,350	36%
$288,350 and above	39.6%

AMT Taxable Income (MFJ)	**Tax Rates**
$0 - $105,000	26%
$105,000 - $161,000	32.5%
$161,000 - $285,000	35%
$285,000 and above	28%

From 1993-2001, the marginal tax rates in the regular income tax schedule increase from 15 percent in the lowest bracket to 39.6 percent in the highest bracket. From 2002 to 2011, marginal tax rates range from 10 percent in the lowest bracket to 35 percent in the highest bracket. In comparison, The AMT has a non-graduated schedule in terms of effective marginal tax rates. In 1993, the Omnibus Budget Reconciliation Act (OBRA) altered the AMT schedule by eliminating a flat marginal tax rate of 24 percent and introducing a two-tiered schedule, with statutory tax rates of 26 percent and 28 percent. The AMT also provided a fixed deduction of $45,000 to married joint filers and $33,750 to single filers. These exemption amounts are phased-out at higher taxable

[11] The detailed legislative history of the AMT is provided in Annex A.

income levels. For example, in year 2000, this phaseout begins at $105,000 for married joint filers and $78,750 for single filers. In the phaseout range, every additional dollar of taxable income reduces the fixed deduction by 25 cents leading to effective marginal tax rates that are 1.25 times the statutory marginal tax rates. The fixed deduction completely phases out at taxable income of $285,000 ($213,750) for married joint filers (single filers) in year 2000, creating an effective AMT schedule consisting of four distinct marginal tax rates: 26 percent, 32.5 percent at the point where the exemption phaseout begins, 35 percent where the 28 percent statutory rate begins and exemption phaseout continues, and 28 percent, at the point where the fixed deduction is completely phased out.

Third, taxable income is defined differently on the two schedules. The regular income tax schedule does not tax all earned income. Instead, it allows taxpayers to subtract certain deductible consumption and excludable income items from their total earned income for taxation purposes. The residual income forms the tax base on which prevailing tax rates are applied. While a discussion of all the exemptions is beyond the scope of this paper, some of the excluded income items include portions of retirement income, certain types of scholarship income, interest gained from municipal bonds and charitable donations received. As compared to excluded income, deductible consumption expenses that favor certain uses of a taxpayer's income include charitable contributions, state and local taxes paid, real estate taxes paid, interest paid on home mortgage, medical expenses, business expenses and miscellaneous expenditure. High-income taxpayers disproportionately use these excludable income and deductible consumption items that are subject to favorable tax treatment. For example, in fiscal year 2010, taxpayers with incomes below $50,000 used 8.8 percent of all medical deductions, 1.4 percent of all state and local tax deductions, and 2.8 percent of mortgage interest deductions. Compare these utilization rates to those of taxpayers with incomes above $100,000, for whom the shares of these deduction amounts were 49.3 percent, 85.6 percent, and 78.3 percent, respectively.[12] The regular income tax code also provides a fixed standard deduction that can be used by taxpayers for whom the above deduction amounts are less than the standard deduction. Prior to the Tax Cuts and Jobs Act (TCJA) 2017, the regular tax schedule also allowed for personal exemptions for each member of the family.

On the other hand, the AMT disallows major deductions such as personal exemptions, the standard deduction, and important itemized deductions such as the state and local tax (SALT)

[12] Estimates computed using the Joint Committee on Taxation's (JCT) "Estimates of Federal Tax Expenditures for Fiscal Years 2011-2015".

deduction, and miscellaneous deductions used primarily by business owners.[13] By redefining taxable income, the AMT causes a larger part of earned income to be counted as taxable that is otherwise sheltered from taxation on the regular schedule. However, the AMT provides a substantial fixed deduction that keeps low- to middle-income taxpayers out of the AMT.

In Panel A of Figure 2, I illustrate the regular income tax schedule using tax rules prevailing in year 2000. Marginal tax rates increase at each of the kinks in the schedule, represented by the change in slope at the kink points. For example, the marginal tax rate in the lowest taxable income bracket is 15 percent, while the marginal tax rate in the highest bracket is 39.6 percent. The length of each interval between kink points depends on the size of income tax brackets. Marginal tax rates and the length of income tax brackets are fully contingent on overarching tax law, and common to all taxpayers. In contrast, the starting point of the tax schedule along pre-tax income represented by the x-intercept is determined by the total amount of allowable regular income tax deductions that a taxpayer claims and therefore, this parameter varies across taxpayers.

Panel B of Figure 2 provides a similar representation for the AMT schedule with corresponding tax brackets and effective marginal tax rates. The x-intercept of the AMT schedule is equal to the sum of the fixed AMT deduction and the regular tax deductions allowed by the AMT. At higher income levels, deductions under the regular tax schedule are on average lower than those under the AMT, by design. Therefore, Figure 2 relates to a high-income/high-deduction type taxpayer with $x\ intercept_{AMT} < x\ intercept_{regular}$. Note that on average, $x\ intercept_{AMT} > x\ intercept_{regular}$ for a low-income/low-deduction type taxpayer. The differences across taxpayers in the amount of deductions taken on the regular income tax and AMT schedules generates variation in the location of the point at which the two tax schedules intersect. This variation is the key reason for the misalignment of the top kinks on the combined schedule and the regular income tax schedule. Further, as I show in Section IV, this variation in the location of the intersection kink disentangles variation in marginal tax rates from variation in taxable income, severing the link between taxable income and unobserved heterogeneity and mitigating a key endogeneity concern associated with the use of bunching estimators on fixed kink points.

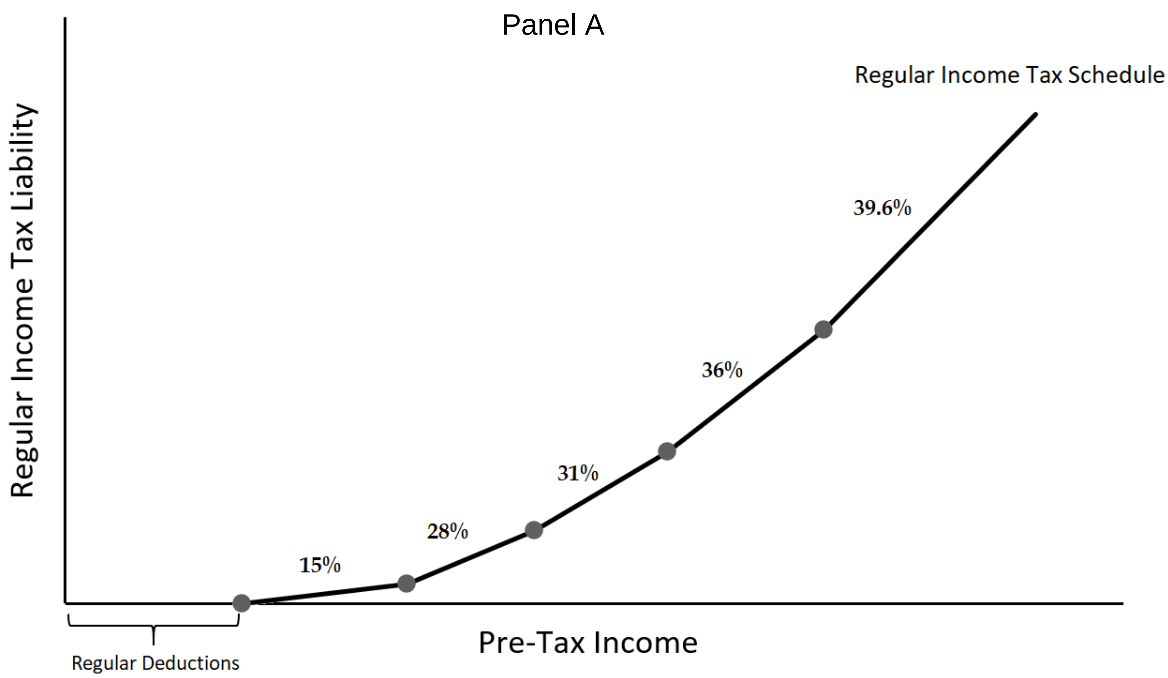
Figure 3, Panel A brings together the regular income tax schedule with the AMT schedule for a high-income/high-deduction type taxpayer.[14] Taxpayers pay the higher of the two personal

[13] The AMT also partially disallows medical and dental deductions, accelerated depreciation, and deductions on home mortgage interest on non-primary property, among others.
[14] High earnings do not automatically translate into higher deductions. However, high earners disproportionately use larger deduction items such as state and local taxes, mortgage interest deduction and medical deductions (JCT

income taxes and therefore, the combined income tax schedule is the upper envelope of the interaction of the two piecewise linear tax functions. The upper envelope of the combined schedule is shaded in gray. The point at which the AMT and the regular tax schedules interact is the intersection kink of the combined schedule. The case for a low-income/low-deduction type taxpayer is different.

The substantial, fixed deduction provided by the AMT shifts the AMT function to the right of the zero pre-tax income point. The fixed deduction between 1993-2011 is as low as $45,000 and as high as $74,450 for married joint filers. This ensures that low- and middle- income taxpayers are only subjected to the regular income tax schedule. In general, this holds true if allowable deductions under the regular tax schedule are generally less than the fixed deduction provided by the AMT. Such a scenario for a hypothetical low-earner/low-deduction type taxpayer is illustrated in Panel B of Figure 3. For these taxpayers, the regular income tax schedule continues to be the effective tax schedule. This is a potential reason for other studies detecting bunching responses for low-income taxpayers when using the regular income tax schedule, but not for high earners who are in fact, subject to the combined schedule. In this paper, I focus on the high-income/high-deduction type of taxpayer responding to the combined schedule in Panel A of Figure 3 to estimate the ETI of high earners in the United States.

Estimates, 2011-2015), which are fully or partially offset by the AMT, leading to high earners having lower deductions on the AMT schedule relative to the regular income tax schedule, on average.



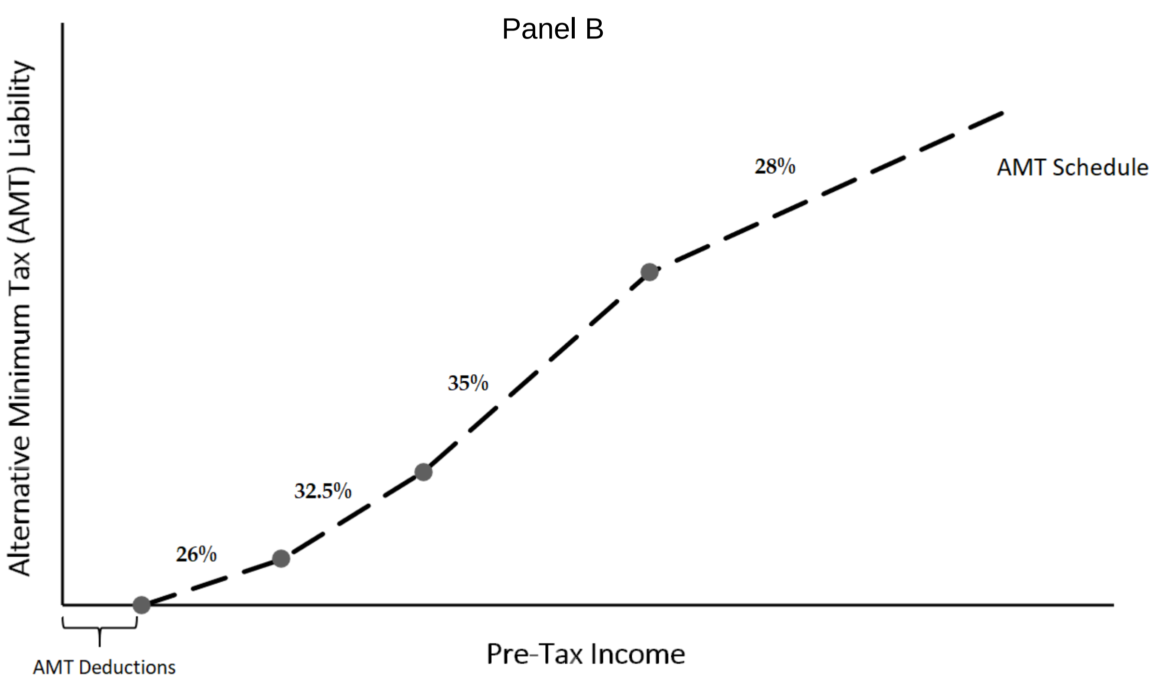


Figure 2: The Regular Income Tax and the AMT Schedules

Notes: Panel A illustrates the regular, federal income tax schedule for a hypothetical taxpayer. The slopes and the length of line segments in the piecewise linear function are based on marginal tax rates and the size of income tax brackets as provided in the tax code for year 2000. The x-intercept is determined by the amount of allowable deductions claimed by the taxpayer under the regular schedule. Panel B illustrates the AMT schedule. This piecewise function corresponds to marginal tax rates and income tax brackes on the AMT schedule. The x-intercept is determined by the amount of allowable deductions claimed under the AMT.The figures are not drawn to scale.

Panel A

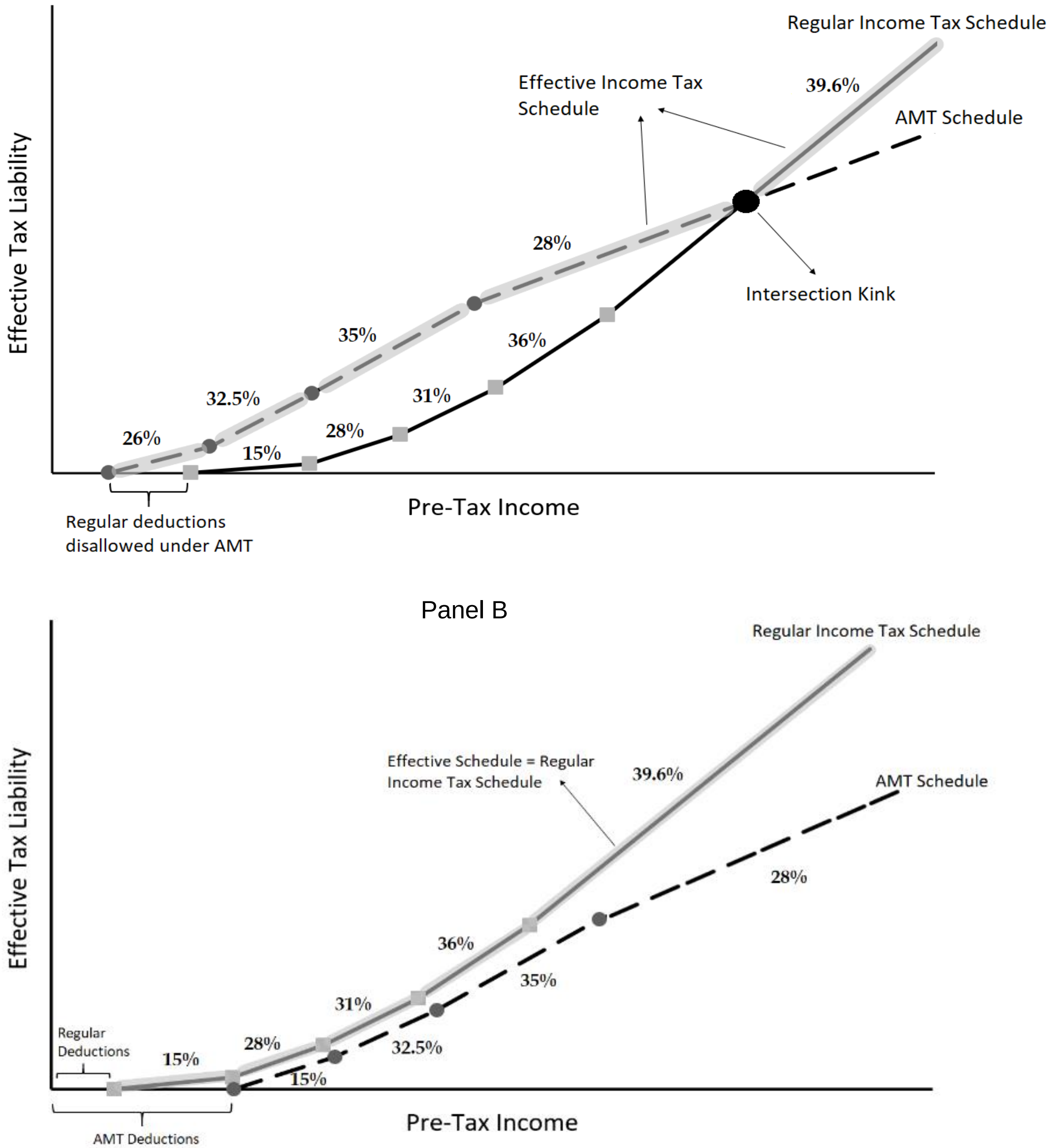


Figure 3: The Combined Schedule

Notes: Panel A is a representation of the combined schedule for a hypothetical high-income/high-deduction type taxpayer. Taxpayers pay the higher of the two schedules, leading to the effective schedule being the upper envelope of the combined schedule, represented in grey highlighting. Panel B illustrates the combined schedule for a hypothetical low-income/low-deduction type taxpayer. For such a taxpayer, deductions on the AMT are, on average, greater than deductions on the regular tax schedule, leading to the AMT function being shifted further to the right, relative to the regular schedule. Since taxpayers pay the higher of the two taxes, the regular tax schedule continues to be the effective schedule for such a taxpayer.

Variation in the location of the top kink in the combined schedule is driven by variation in the difference in the x-intercepts of the two schedules, with the latter depending on the difference in the amount of deductions allowed under the regular tax and AMT schedules. Specifically, one could imagine a range of differences in the x-intercepts, only one of which is illustrated in Figure 3, Panel A, generating a range of intersection kinks. In Figure 4, I provide the distribution of intersection kinks along regular taxable income to in the data to illustrate the variation in the location of the intersection kinks. The figure disaggregates the overall, bimodal distribution into two separate distributions corresponding to time periods 1993-2002 and 2003-2012. Changes made to the tax code through increases in AMT fixed deduction amounts from 2003 onwards

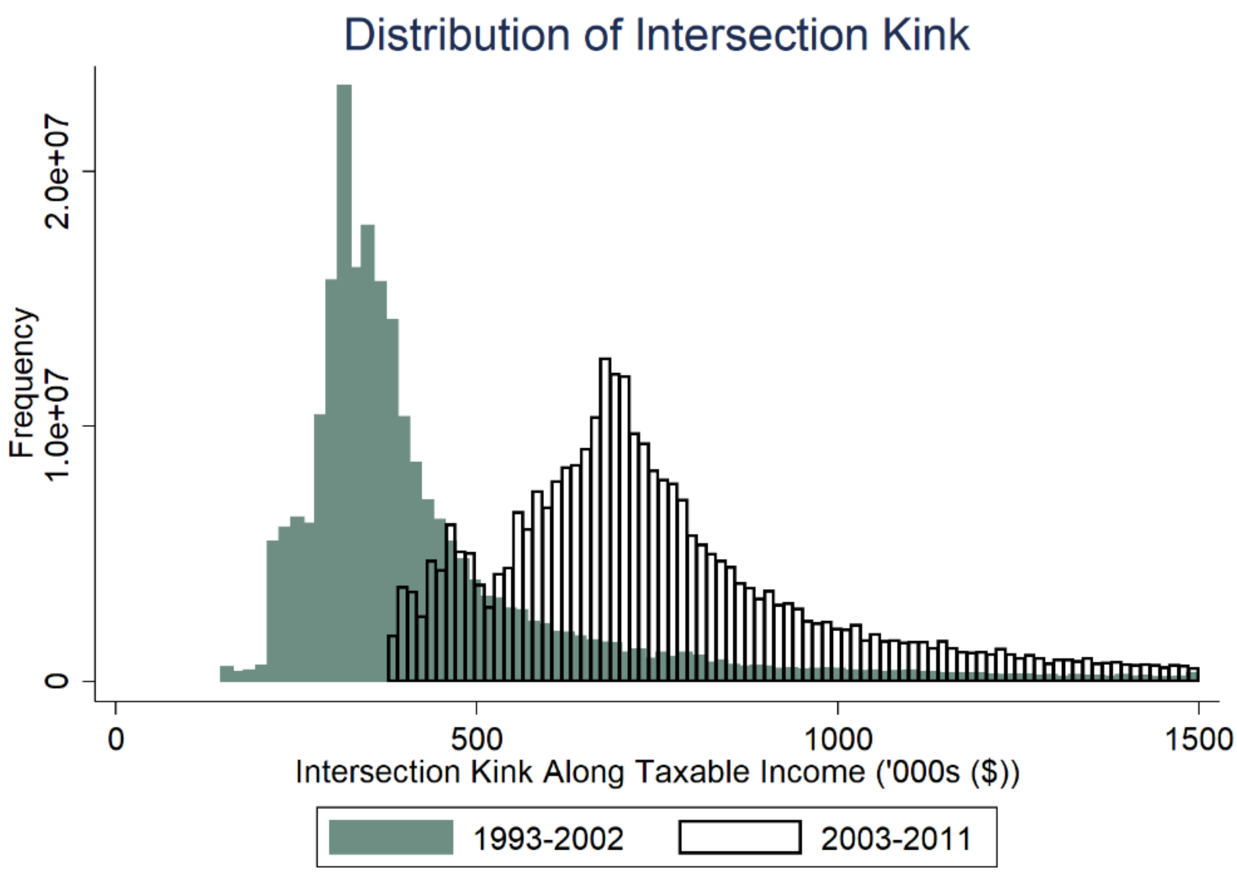


Figure 4: Distribution of the Intersection Kink's Location Relative to Regular Taxable Income

Notes: The location of the itnersection kink in the combined schedule varies across taxable income, unlike kinks on the individual schedules that are fixed in taxable income. The bimodal distribution for its location relative to regular taxable income is divided in two, with the shaded distribution representing time period 1993-2002, and the unshaded distribution representing time period 2003-2011. Tax reforms of 2003 followed by annual increases in AMT exemption amounts shifted the underlying AMT schedule to the right, leading to the intersection kink also shifting to the right in the combined schedule. This leads the distribution for 2003-2011 to appear at higher taxable income levels.

increasingly shifted the AMT function to the right, shifting intersection kinks, on average, to higher income levels.

In Section IV, I use the features of the regular income tax and AMT schedules discussed here to construct the two tax functions for each taxpayer in my sample from 1993-2011. For each taxpayer in Table 1, who is subject to the combined schedule, I solve the two piecewise linear tax functions to find the top, intersection kink in the combined schedule. I use information on each taxpayer's observed taxable income and the location of the taxpayer-specific intersection kink to assess how far the individual's taxable income lies from the intersection kink. Aggregating across taxpayers, I show evidence of bunching to the left of the intersection kink where the marginal tax rate is 28 percent, as compared to approximately 38 percent on average, to the right of the kink.

IV. Empirical Methodology

Data

Since 1960, the Statistics of Income (SOI) division of the Internal Revenue Service (IRS) has published annual samples of individual tax returns in the form of Public Use Files (PUF). These micro-datasets are generated using a stratified random sample of tax filers. Weights associated with sampling have varied – high earners face a larger sampling rate, with those at the very top of the income distribution facing an approximately 33 percent rate of sampling. Since this study specifically looks at high earners, such oversampling allows me to capture greater variation in tax returns for this subpopulation.

I use this income tax return data from 1993-2011, housed at the National Bureau of Economic Research (NBER). In the year 1993, the Omnibus Budget Reconciliation Act (OBRA) altered the AMT schedule by eliminating the flat marginal tax rate of 24 percent and introducing a two-tiered schedule, with statutory tax rates of 26 percent and 28 percent. OBRA 1993 also introduced a fixed deduction of $45,000 on alternative minimum taxable income for married joint filers and $33,750 for single filers. As discussed in Section III, the phaseout of the fixed deduction created four effective marginal tax rates: 26 percent, 32.5 percent, 35 percent, and 28 percent. The IRS stopped providing raw data in 2012, beyond which data available at the NBER contains inflated figures based on 2012 tax return information. Further, the American Taxpayer Relief Act (ATRA) of 2012 indexed the AMT exemption amounts to inflation. To avoid this tax year with characteristics that are significantly different from those found in other tax years, I omit the year from the analysis.

For this analysis, I also divide the sample into two time periods for heterogeneity analysis: 1993-2002 and 2003-2011. I choose 2002 as the endpoint for the first time period because while the AMT fixed deduction amounts were not indexed for inflation before 2012, Congress increased the deduction amounts annually on an ad-hoc basis from 2003 onwards. These increases shifted the AMT schedule to the right along the range of pre-tax incomes, leading to the intersection kink appearing at higher income levels.

I limit the data to tax returns submitted by married joint filers and single filers, leading to a dataset containing 2.3 million observations, representing approximately 2 billion unique tax returns. 5.3 percent of taxpayers in the data submit Form 6251, the form used to compute AMT liability (the unweighted fraction in the data sample is 34 percent). However, this fraction increases to 24 percent for taxpayers with adjusted gross income (AGI) in real 2007 terms greater than $100,000, and to 58.5 percent for taxpayers with real AGI greater than $200,000. The IRS puts the burden of submitting the AMT form on the taxpayer. This implies that in case Form 6251 is not submitted and the IRS predicts that the taxpayer would owe AMT liability, then there is possibility of audit. Between 2006 to 2011, taxpayers also had access to an IRS-provided web tool called the AMT Assistant, which required responses to a handful of questions related to the income level and filing status of the taxpayer to recommend the submission of Form 6251.

I limit my analysis to taxpayers who face the combined schedule to assess behavioral response around the intersection kink. This leads to an analysis sample containing 273,856 observations representing approximately 5.9 million tax returns. Further, in line with earlier literature, I restrict the frame of the analysis to a range within which the effective kink lies. I limit the sample to individuals within $300,000 (-$150,000, +$150,000) of their effective kink. This is my analysis sample, with a total of 36,639 observations, representing approximately 1.2 million individual tax returns.

The population median AGI for these individuals is $621,800 in real 2007 dollars, corresponding to taxpayers in the top percentile of the income distribution. The median, real effective taxable income for taxpayers in the analysis window is $536,601. The intersection kink for these taxpayers lies on average, at $441,127 for time period 1993-2002, and at $678,706 for time period 2003-2011.

Locating the Top Kink in the Combined Schedule

This section details the methodology for constructing the combined schedule as the combination of the regular income tax and AMT schedules, overlaying the distribution of taxable income across the combined schedule, and leveraging bunching around the intersection kink point – the point where the AMT and the regular income tax schedules intersect – to estimate the elasticity of taxable income with respect to the net-of-tax rate. As discussed, a taxpayer can shelter part of his or her pretax income from taxation by taking deductions under both the regular income tax and the AMT. Let the pre-tax income in a calendar year for a given taxpayer be Y. Let taxable income sheltered under the regular income tax schedule via deductions be D_R.[15]

The AMT has a fixed deduction for each tax return filing category that I denote by D_{AMT}. Further, the AMT disallows part of the deductions taken under the regular income tax. Therefore, part of the deductions taken under the regular income tax schedule which are partially allowed under the AMT are $D_R(1-\alpha)$, where α is the proportion of regular income tax deductions disallowed under the AMT. If tax sheltering (S_R and S_{AMT}) under both schedules is equal:

$$S_R = S_{AMT}\text{, then } D_R = D_{AMT} + D_R(1-\alpha)\text{, or } \alpha = \frac{D_{AMT}}{D_R}$$

On average, if $S_{AMT} < S_R$, as is the case in Figure 3, Panel A, then the taxpayer has a unique intersection kink. If $S_{AMT} \geq S_R$, the taxpayer's effective tax schedule can either have two intersection kinks – one at a low level of taxable income and the other at a high level of taxable income, or no intersection kinks, as shown in Panels B and C, respectively. As discussed earlier, to ensure analysis of a unique intersection kink, I restrict the sample to taxpayers for whom $S_{AMT} < S_R$. This leads to the analysis sample containing 273,856 observations representing approximately 5.9 million tax returns.

[15] In fact, overall sheltered income includes "exemptions" and "deductions". For simplicity, I combine both and call them deductions.

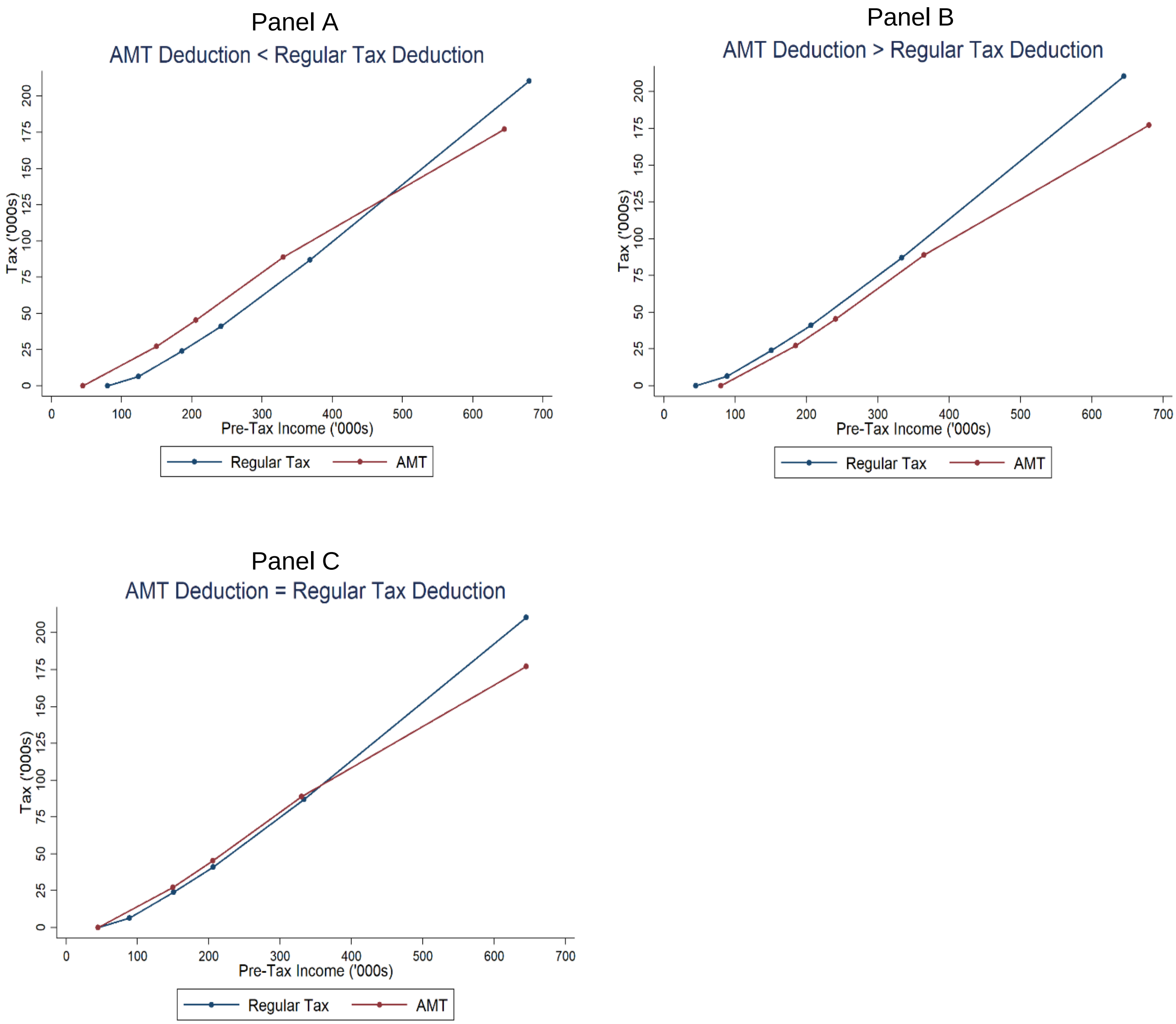


Figure 5: Relationship between Deductions and the Location of the Intersecton Kink

Notes: These figures show how the location of the intersection kink varies when the difference in deductions allowed between the regular income tax and the AMT schedules differs. The current study isolates the analysis to taxpayers who exhibit a structure of deductions and the combined schedule like the one shown in Panel A.

The two schedules are piecewise linear. To find the intersection kink, I solve the system of two piecewise linear functions to find the location of the intersection kink. I take the AMT schedule as given, starting at the zero taxable income point. However, to adjust for the distance between the starting points of the two schedules, I treat the difference in deductions as an x-intercept dummy. Since the IRS PUF data does not provide exhaustive information on deductions, I exploit the difference between the alternative minimum taxable income and the regular taxable income. This precisely equals the difference in deductions allowed under the two schedules. Let T_R be regular taxable income, defined as $T_R = Y - S_R$. Let T_{AMT} be alternative minimum taxable income, defined as $T_{AMT} = Y - S_{AMT}$.

Then,

$$T_{AMT} - T_R = (Y - S_{AMT}) - (Y - S_R) = S_{AMT} - S_R$$

As I discuss earlier, the location of the intersection kink varies across taxpayers, so I standardize the location of this kink by centering the distribution of taxable income relative to the intersection kink. I do this by finding the distance of effective taxable income for each taxpayer from the intersection kink on his or her respective effective income tax schedule, and then plot the distribution of these differences relative to the intersection kink. To visualize this distribution, I construct histograms with taxable income bins of varying widths. The choice of binwidth leads to a trade-off between noise and precision: the greater the binwidth, the less noisy and smoother the histogram; the smaller the binwidth, the noisier the histogram, since it reveals more variation in the data. I compute the optimal binwidth using a data-driven approach. I also use other binwidths for comparison. For the optimal bindwidth selection, I use the Freedman-Diaconis method:

$$k = 2 * IQR * n^{-1/3}$$

Where k is the binwidth, IQR is the interquantile range of the distribution of effective taxable income, and n is the number of observations. I find k to be \$8,105.

Estimating the Elasticity of Taxable Income

To estimate the elasticity of taxable income, I use the traditional bunching estimator developed by Saez (2010). This will be the base estimator for my analysis, on which I impose bounds by utilizing weaker assumptions, as discussed below. For the base estimator, I exploit a simple parameterized model with a quasi-linear and iso-elastic utility function of the form:

$$u(c,k) = c - \frac{n}{1+1/e}\left(\frac{k}{n}\right)^{1+1/e}$$

Where c is consumption, k is before-tax income, n is an ability parameter distributed with density $f(n)$, and e is compensated elasticity of reported income (Saez, 2010). By using a quasi-linear utility function, I abstract from any income effects for simplicity. Maximizing this with a linear budget constraint and letting $L_0(k)$ be the cumulative distribution of earnings when there is a constant marginal tax rate τ_0 throughout the distribution, the introduction of a convex kink in the budget set at K introduces the new tax rate τ_1 that is applied to pre-tax income which is over and above K. Taking this kink point into account, individuals with $n \in [K/(1-\tau_0)^e, K/(1-\tau_1)^e]$ choose $k = K$ and bunch at the kink point. This leads to the fraction of the population bunching:

$$b = K\left[\left(\frac{1-\tau_0}{1-\tau_1}\right)^e - 1\right]\frac{l(K)_- + l(K)_+/\left(\frac{1-\tau_0}{1-\tau_1}\right)^e}{2}$$

Which can be solved explicitly to express e as a function of observable or empirically estimable variables. Simplification leads to:

$$\epsilon = \frac{b(\tau_0, \tau_1)}{K\log\left(\frac{1-\tau_0}{1-\tau_1}\right)} \approx \frac{\hat{b}}{\left|\frac{K}{W}\cdot\frac{\Delta\tau}{1-\tau_0}\right|}$$

Where τ_0 and τ_1 are the effective marginal tax rates on either side of the intersection kink and are observed. For example, in year 2000, τ_0 is 28 percent and τ_1 is 39.6 percent. W is the binwidth chosen for binning taxpayers in effective income groups. The traditional bunching estimator uses a fixed K in taxable income. However, since the location of the intersection kink varies along the regular taxable income spectrum, I take the weighted average of the effective taxable income in the bunching region as an estimate of K. $\hat{b}$ measures the difference between the observed taxable income density in the presence of the kink point, and the counterfactual density which would plausibly have existed in the absence of the kink point. In other words, $\hat{b}$ quantifies excess mass, or the magnitude of bunching in the bunching region. Plugging in the observed marginal tax rates and estimates of excess bunching $\hat{b}$ and K, provides base estimates of the elasticity of taxable income.

To estimate b, Saez (2010) assumes the counterfactual density to be linear in the bunching region. However, I fit a polynomial function across the bunching region to estimate the counterfactual density, in the spirit of Chetty et al. (2011). Dividing the range of taxable incomes relative to the intersection kink into bins given sizes, I fit a polynomial of order 7 to the counts for each of the taxable income bins, excluding data near the kinks by estimating a regression of the form:

$$C_j = \sum_{i=0}^{p} \beta_i Z_j^i + \sum_{-r}^{r} \phi_k D_j + \epsilon_j$$

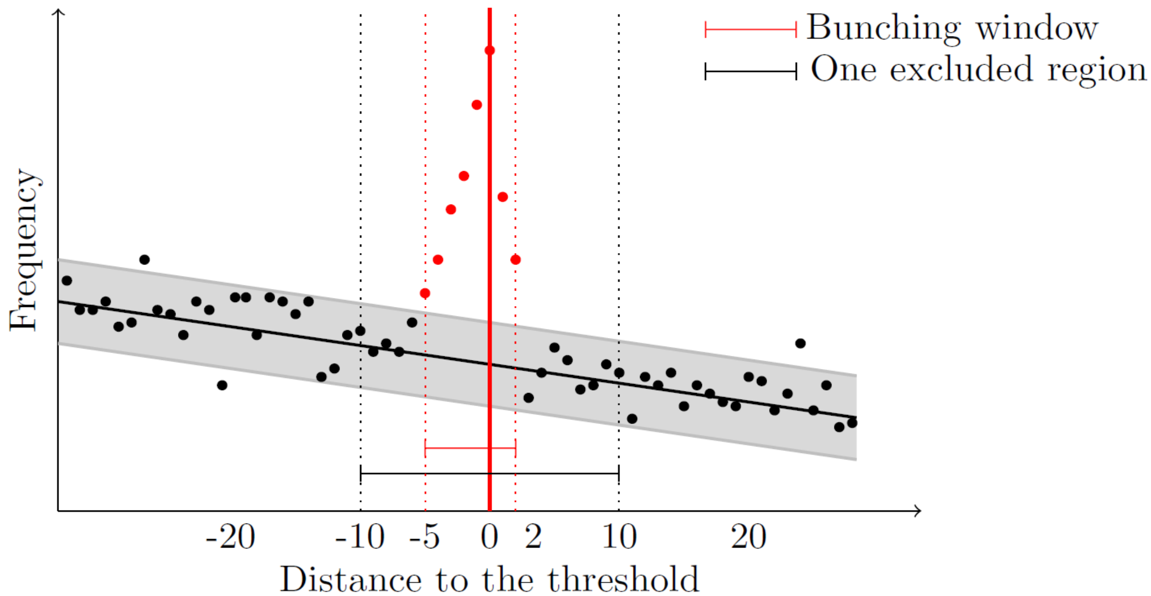
Where C_j is the count of observations found in bin j, Z_j is the midpoint level of the effective taxable income in bin j, and D_j is a dummy for each bin found in the bunching region. In other words, there are r indicators such that $D_j = 1$ if $Z_j \in [K - e, K + f]$, where K is the location of the kink and e is the distance to the left of the kink and f is the distance to the right of the kink measured in terms of effective taxable income. The counterfactual frequency of observations, $\hat{C}_j^{cf}$, is then derived using predicted counts from $\hat{C}_j^{cf} = \sum_{i=0}^{p} \hat{\beta}_i Z_j^i$, which omits the impact of the dummies. Using the actual and the estimated counterfactual densities, the quantity of "excess bunching" can be estimated using:

$$\hat{b} = \sum_{j=c+l}^{c} \frac{(C_j - \hat{C}_j^{cf})}{N}$$

I further impose the constraint that taxpayers who bunch do so by reducing their taxable income, so that the number of taxpayers missing from the right of the intersection kink is equivalent to the number of individuals bunching to the left of the intersection kink. I calculate the standard error for $\hat{b}$ using a parametric bootstrap procedure, by drawing from the estimated vector of errors for the counterfactual estimation equation with replacement to generate a new set of counts, and applying the above technique to calculate a new estimate of $\hat{b}^k$. I define the standard error of $\hat{b}$ as the standard deviation of the distribution of $\hat{b}^k s$.

I construct the bunching region using the algorithm for bandwidth selection proposed by Bosch et al. (2020). The construction of the bunching region comprises two choices: the choice for the location of the bunching region (symmetric or asymmetric) and the length of the bunching region on either side of the kink point. Earlier methods for selection of this model parameter in the literature have been ad hoc, depending primarily on either visualizing bunching and selecting bandwidth to capture this observed bunching, or to have a symmetric bandwidth with sensitivity analysis assessing changes in elasticity with changes in bandwidth size.

The algorithm for selecting the bandwidth is as follows. The bin with its midpoint at the kink is initially considered to be the bunching region, so that the excluded region becomes $(z_{-}, z_{+}) = (0,0)$. A local linear regression is run using all observed bins outside the excluded region defined as the kink point, leading to predicted frequencies in each bin and the 95 percent confidence interval. Observed, contiguous bins outside the confidence interval are taken to be the bunching region, with the left-most bin being the lower bound of the bunching region, and the right-most bin being the upper bound. The excluded region is then recursively increased to cover all other excluded regions, with a bin being added to the left and the right each time, so that $z_{-} \in \{-Z, (-Z+1), (-Z+2), \ldots, 0\}$ and $z_{+} \in \{0, 1, \ldots, Z\}$ and the same local linear regression method implemented to obtain distributions for the lower and upper bounds of the excluded region. For an illustration of this method, see Figure 6. The mode of the lower (upper) bound distribution forms the lower (upper) bound of the bunching region for my core analysis. These correspond to \$-44,059 and \$+16,212 in my analysis sample. While I utilize the algorithm for selecting the bunching window below, I also conduct a test for sensitivity to varying choices of bandwidth in Section VI.



Source: Bosch et al. (ITPF, 2020)

Figure 6: Data-driven Selection Algorithm for Location and Size of Bunching Region

Notes: This figure illustrates the algorithm for selecting the bunching region. Each point on the figure represents the frequency of taxpayers being in each taxable income bin. The solid, vertical line identifies the location of the kink point. The straight line running through it is the line of best fit when a local linear regression is fit through the observed bins except those that are excluded, and the shaded area is its confidence interval. The excluded region is incrementally expanded from (0,0) to (-X,+X), providing distributions of the upper and lower bounds of the bunching region. The modes of these distributions are selected as the starting and ending points of the bunching window.

I further address such endogeneity concerns by imposing lower and upper bounds on the average elasticity estimate for high earners. I exploit the method developed by Bertanha, MacCallum and Seegert (2016, 2020) and assume that the counterfactual density belongs to the family of Lipschitz continuous functions. In the context of the counterfactual density assumed to be defined by a Lipschitz function, there exists a real number such that the line connecting the endpoints of a given bunching region has a slope which is not greater than the absolute value of this real number, known as the Lipschitz constant. This limits the magnitude of bunching (or trough) of the underlying density. Such a limitation is achieved by constraining the area under the counterfactual density with the area of the observed distribution[16]. With these limits, observed

[16] I thank Nathan Seegert at the University of Utah for providing me with early access to their statistical program for identifying these bounds. The final Stata package is available under the label “bunching”.

density can be compared with the bounded counterfactual density to generate bounds of elasticity estimates for high earners. Results for this robustness check are provided in Section VI.

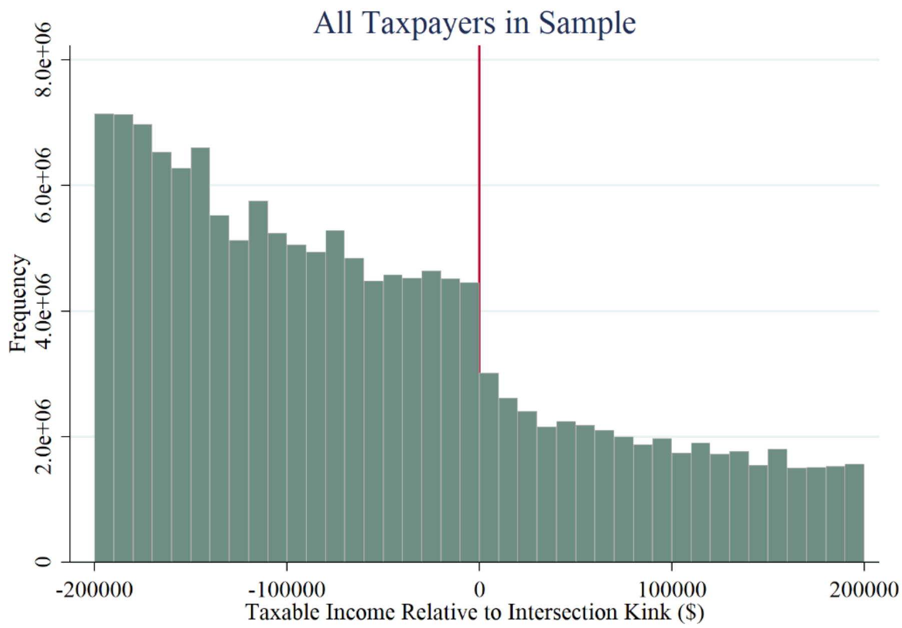
V. Discussion of Results

Graphical Evidence

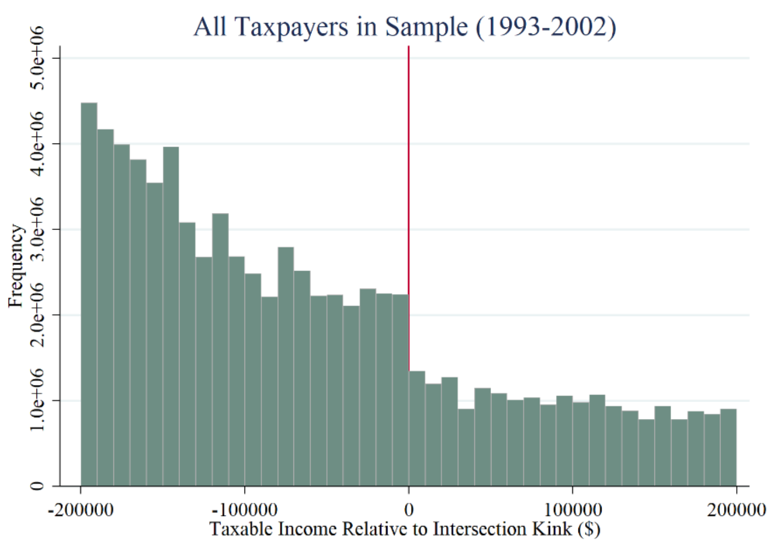
The above algorithm reveals graphical evidence of clustering to the left of the intersection kink, as shown in Figure 7, Panel A. This figure provides the weighted distribution of taxable income relative to the intersection kink for taxpayers in the sample. Note that to give an expanded view of the distribution around the intersection kink point, I plot the observed distribution within -$200,000 to +$200,000 of the intersection kink. Panels B and C provide histograms disaggregated by the time periods 1993-2002 and 2003-2011. Both periods reveal bunching response just to the left of the intersection kink, with more pronounced bunching for the latter period. With increasing AMT bunching amounts, the intersection kink shifts towards the right along the taxable income distribution. The accentuated bunching response revealed in the time period 2003-2011 arguably captures the potentially higher behavioral response at relatively higher income levels.

Figure 8 contrasts the bunching responses to the intersection kink and around top kinks in the regular income tax schedule. Panels B and C are come from Saez (2002). To ensure visual comparability, I also provide the histogram for the entire time period in my analysis with smaller binwidths of $4,000. Given the higher income level at individual intersection kinks, the observed distribution around this kink point leverages a thinner part of the income distribution leading to fewer observations. To increase variation in the observed distribution, I utilize my entire time period of 1993-2011, as compared to the time period 1993-1997 considered in Saez (2002). This comparison is, therefore, only illustrative of the evidence for non-trivial bunching around the intersection kink relative to individual kinks on the regular income tax schedule.

Panel A



Panel B



Panel C

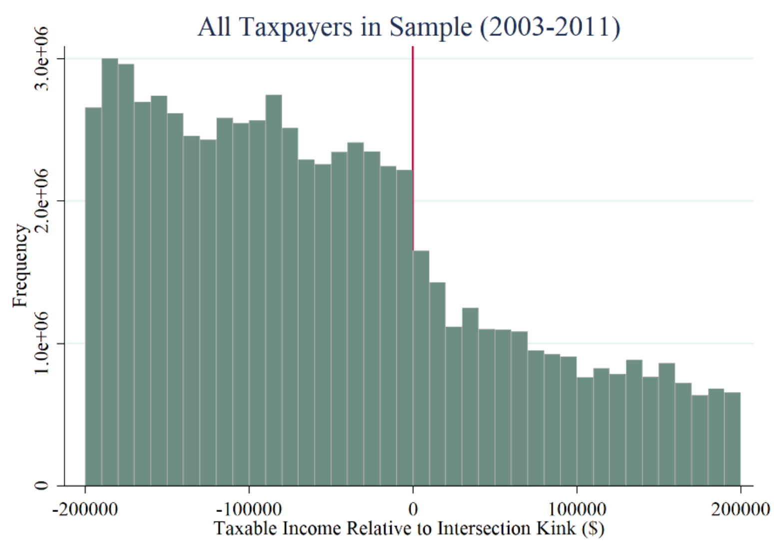


Figure 7: Graphical Evidence of Bunching

Notes: The figures show bunching of taxpayers in the aggregate around the intersection kink. Histograms have binwidth of $10,000. Panel A shows the distribution of effective taxable income relative to the intersection kink for all observations in the study sample. Panels B and C show the distributions for subpopulations disaggregated by two time periods: 1993-2002 and 2003-2011.

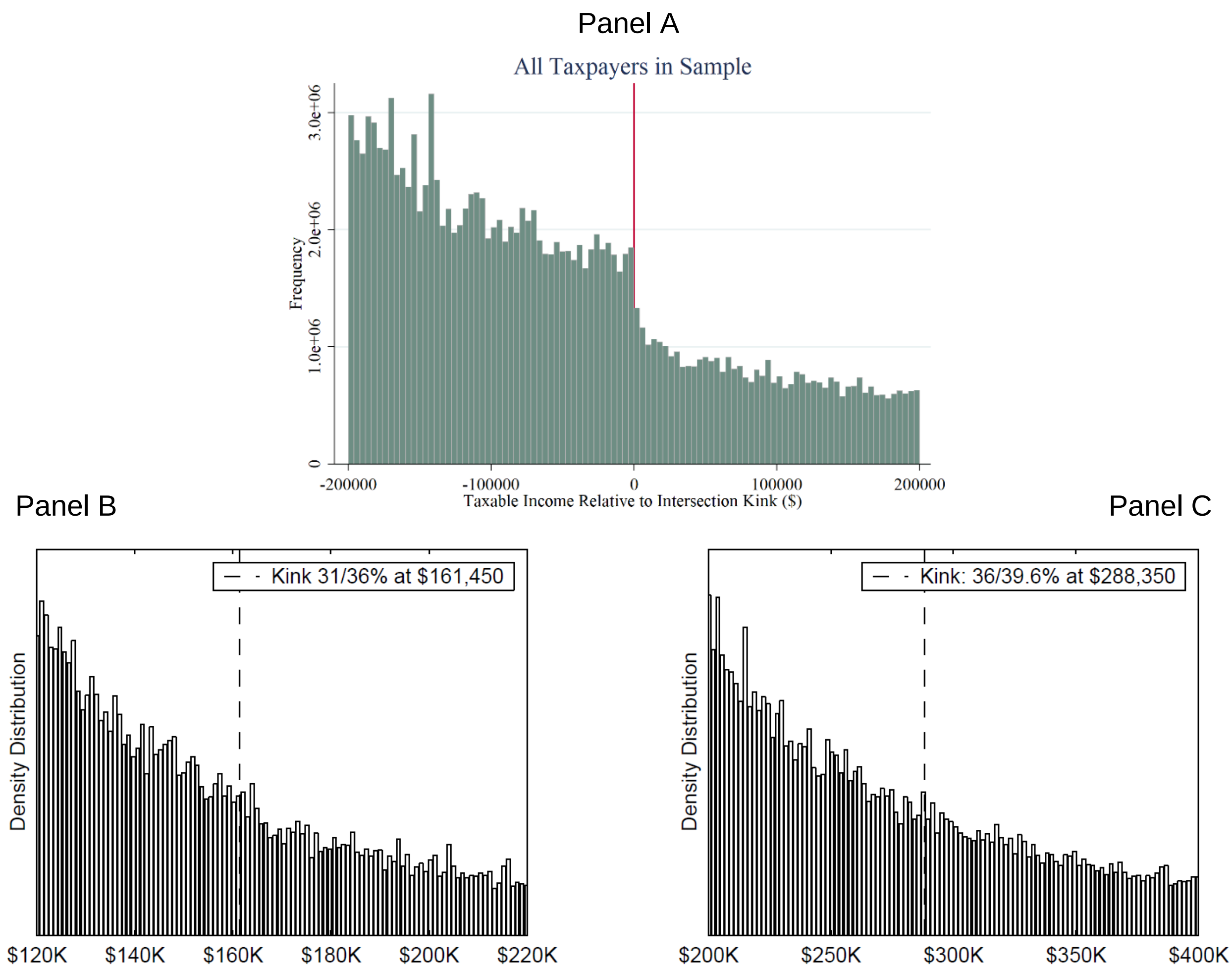


Figure 8: Bunching Response around the Top Kinks of the Combined and the Regular Schedules

Notes: The figures show bunching responses around the intersection kink in the combined schedule (Panel A), and the top two kinks in the regular income tax schedule when the combined schedule is not considered (Panels B and C). Histograms have binwidth of $4,000. Panels B and C are sourced from Saez (2010).

I further disaggregate the total sample into wage earners and the self-employed. Self-employed are defined as taxpayers who revealed any non-zero income from non-wage sources, including sole proprietorship; partnerships and S-Corporations; and farming. Wage earners are those taxpayers who reported zero earnings from these sources. I will refer to taxpayers with any positive self-employment earnings as self-employed, though this does not preclude them having wage-based income as well. Existing literature has predicted, and shown for other segments of the income distribution, significant avoidance behavior by self-employed individuals as compared to wage earners. Pure wage earners in the United States face third-party reporting, with their employers sending the W-2 form containing information on the employees' earnings to the IRS, which the IRS uses to test for any mismatches between employee- and employer-reported incomes, increasing the probability of detection of tax avoidance or evasion for pure wage earners. Self-employed individuals face third-party reporting for only a fraction of their overall income which comes from wages, and therefore, have greater flexibility in reporting their self-employment income, providing them with a larger margin to manipulate taxable income.

Graphical evidence in Figure 9, Panel A shows that high earning pure wage earners also cluster to the left of the intersection kink, though relatively less sharply as compared to taxpayers who have access to self-employment income (Panel B). The substantial bunching for high-income wage earners is in contrast to earlier studies which show very low bunching for wage earners in the overall taxable income distribution (Saez, 2010; Chetty et al. 2011). Two reasons possibly lead to this divergence. First, high-income wage earners have increased bargaining power, allowing them to negotiate the substitution of highly taxed monetary compensation with untaxed fringe benefits. Second, high-income managers and executives get a larger share of their earnings in the form of stocks as compared to lower-income wage earners. The realization of gains or losses on such stocks can be timed flexibly relative to annual wage earnings.

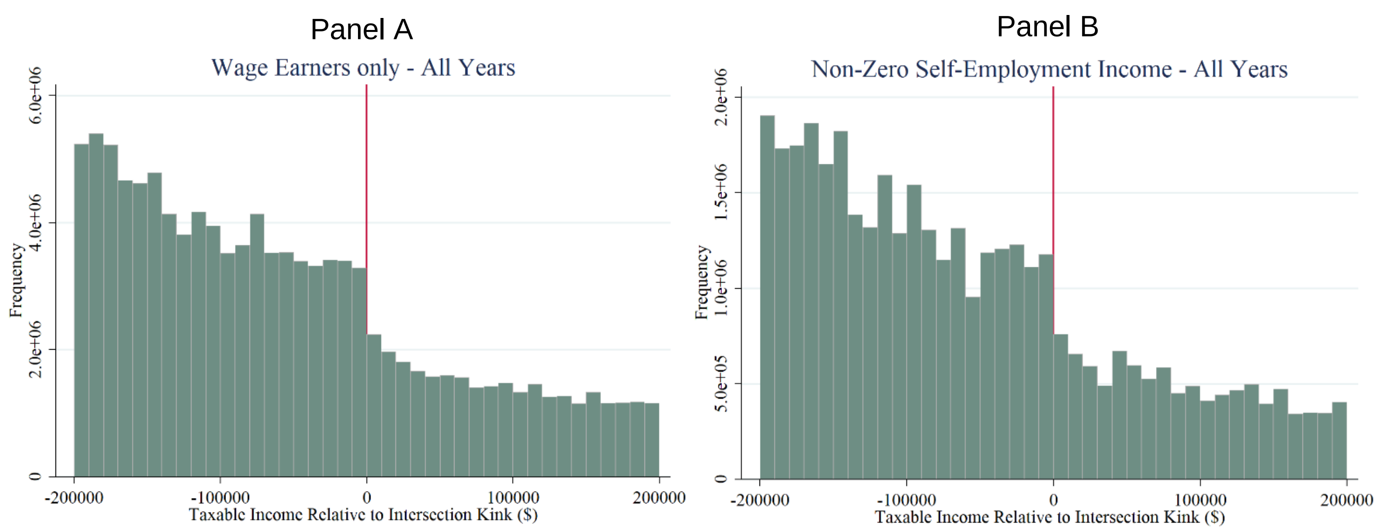


Figure 9: Bunching Responses of Wage Earners and the Self-Employed

Notes: These figures provide visual evidence of bunching around the intersection kink for pure wage earners (Panel A) who do not report any self-employment income, and for taxpayers with any positive self-employment earnings (Panel B). Histograms have binwidth of $10,000.

The availability of capital stock, and strategic realizations of capital gains and losses provide high-income taxpayers with the ability to optimize tax sheltering. Note that for the time period considered in this paper, short-term capital gains are taxed at the same rates as ordinary income and therefore, divergent strategies for tax sheltering using long-term capital gains are unlikely. However, if long-term capital gain or loss realizations are timed strategically, then there is potential to use capital stock activity to shed light on tax avoidance mechanisms beyond the AMT.[17] Such avoidance behavior, especially where realizations are timed according to current and future expected tax rates, can give rise to fiscal externalities which also need to be incorporated into elasticity estimates which can then be used to estimate efficiency costs of income taxation under certain conditions.

Assessing strategic behavior on the capital gains channel, however, is difficult for two reasons. First, the cross-sectional nature of publicly available IRS tax return microdata is not amenable to assessing timing of realizations for the same taxpayers. Utilizing the universe of tax data at the IRS can allow for more flexibly in studying these mechanisms. Second, in the context of the interaction of the regular income tax schedule and the AMT, assessing the impact of capital gains implies overlaying a third schedule on top of the first two schedules. To avoid this complexity, I divide the sample into individuals reporting long-term capital gains and those who

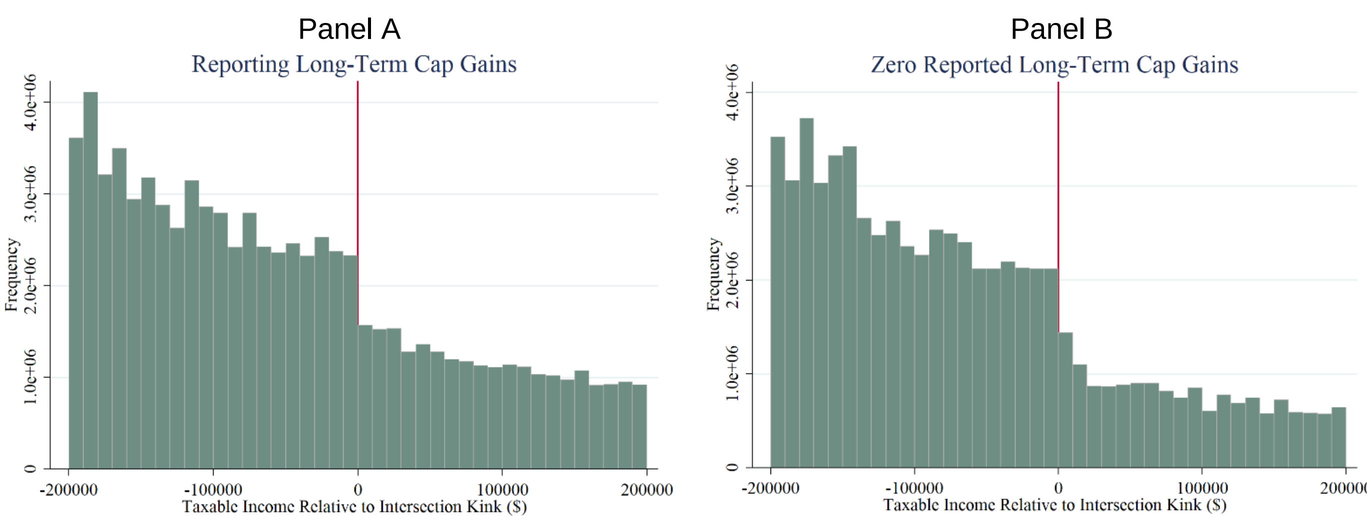


Figure 10: Bunching Responses of Capital Gainers and Capital Non-Gainers

Notes: These figures provide visual evidence of bunching around the intersection kink for taxpayers reporting long-term capital gains (Panel A) and those not reporting such gains (Panel B). Histograms have binwidth of $10,000.

[17] Realizations of short-term capital gains do not occur in isolation from strategies for long-term capital gain realization. For simplicity however, I treat all short-term capital gains as ordinary income and abstract from their effect on the ability of taxpayers to realize long-term capital gains.

do not report such gains, and separately find graphical evidence and elasticity estimates for both groups. Thus, estimates for the subpopulation of tax returns only affected by the combined schedule which is a combination of the regular income tax schedule and the AMT, and not the capital gains schedule, provide the clearest insights into taxpayer behavior around the intersection kink. I find that individuals who report no long-term capital gains in a given year have a greater bunching response to the intersection kink, as opposed to taxpayers who do report such gains. These results are illustrated in Figure 10.

Estimates of the ETI for High-Income Taxpayers

Figure 11 illustrates the observed and counterfactual distributions of effective taxable income relative to the kink point for high earners in my sample of tax returns filed between 1993-2011. The blue line represents the observed density, and the smooth red line represents the counterfactual density. The zero point in the support of the distribution represents the recentered location of the intersection kink point. The vertical dashed lines illustrate the bounds of the bunching region. The observed distribution contains some noise on the left tail of the distribution in the analysis window, but it cancels out in the aggregate. The density of observed effective taxable income around the combined schedule intersection kink point provides evidence that high earners by bunching to the left of the intersection kink, with the difference between net-of-tax rates between the left and the right of the intersection kink being approximately 16 percent.

Using this bunching response, I estimate the elasticity of taxable income with respect to the net-of-tax rate for high earners to be 0.15, estimated precisely within a 95 percent confidence interval. This estimate is economically significant as compared to earlier estimates of close to zero for high-earners in the United States, obtained with the use of bunching estimators and the regular income tax schedule (Saez, 2010; Mortenson & Whitten, 2016).

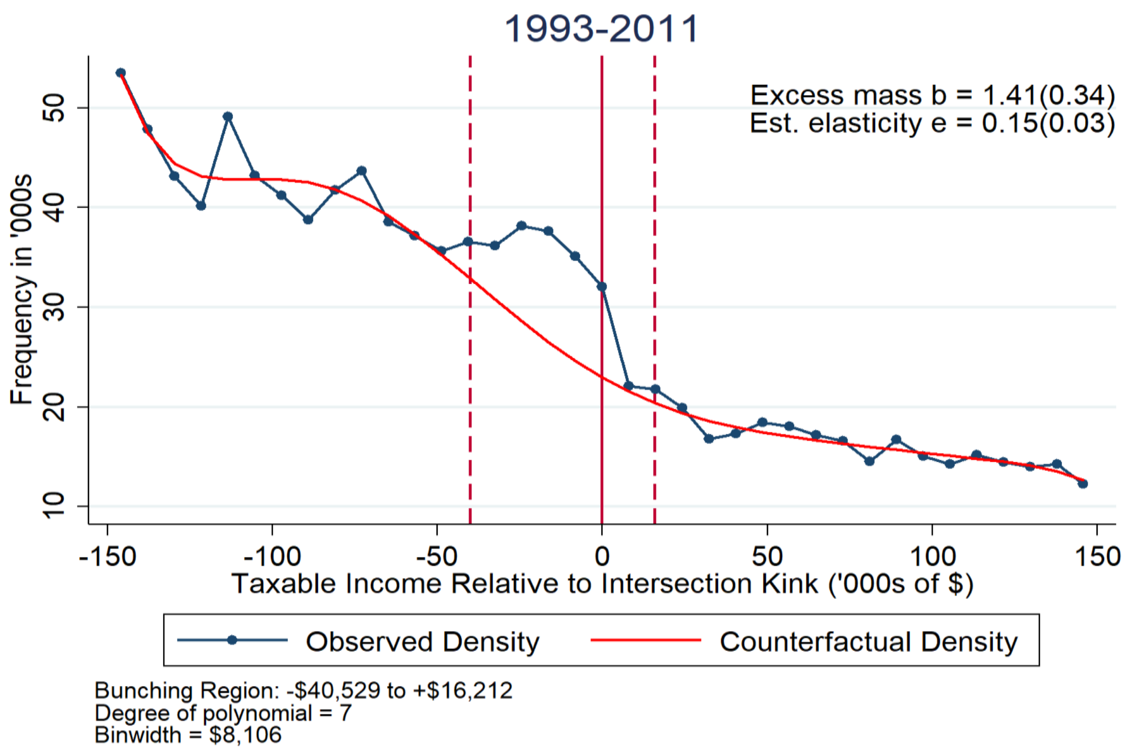


Figure 11: Distribution of Observed Versus Counterfactual Effective Taxable Incomes

Notes: The bold, vertical line (red) represents the centered intersection kink. Dashed, vertical lines (red) represent the lower and upper bounds of the bunching region defined as $-40,529 and $16,212. Observations are binned, with optimal binwidth of $8,106. A seventh-order polynomial is used to construct the counterfactual distribution. Estimates for excess mass and elasticity are provided at the top-right. Bootstrapped SEs are provided in parentheses.

In my analysis sample, 28 percent of taxpayers report at least some non-zero self-employment income. The remaining 72 percent only report pure wage earnings. As I discuss in Sections V, taxpayers with self-employment income do not face third-party reporting for at least some part of their incomes, creating space for tax avoidance behavior that can be more aggressive relative to the behavior of pure wage earners, for whom such tax avoidance space is more limited. However, high-wage employees such as executives and managers might also have greater bargaining power as compared to low-wage employees, vis-à-vis fringe benefits and converting part of wage income into lower-taxed stock options. While previous literature has shown that bunching responses of wage earners for the overall population are weaker, it is plausible that these responses are not trivial for high-income wage earners.

Panels A and B in Figure 12 show that both pure wage earners and taxpayers with some self-employment income bunch to the left of the intersection kink. However, this bunching is more pronounced for taxpayers with self-employment income as compared to taxpayers with only wage earnings. I estimate elasticity of taxable income for wage earners to be 0.12, as compared to estimated elasticity of 0.24 for individuals reporting non-zero self-employment income, twice that of wage earners. The estimated elasticity for the self-employed is remarkably similar to the observed elasticity for this subpopulation in Denmark (Chetty et al., 2011).

I also divide the analysis sample across taxpayers who do not report long-term capital gains and those who do. Approximately 43 percent of the population represented by the analysis sample does not report long-term capital gains, as opposed to 57 that does report some form of long-term capital gains. Panels C and D in Figure 12 provide evidence of bunching to the left of the intersection kink for both groups. However, this bunching response is accentuated for taxpayers not reporting long-term capital gains. It is possible that for the group reporting such gains, the added complexity of the capital gains schedule on top of the interaction of the regular income tax and AMT schedules results in some of the bunching at the kink point being dispersed. Specifically, the capital gains schedule can create a wedge between the combined AMT-regular income tax schedule and the true combined schedule. Therefore, the cleanest elasticity estimate for high earners responding to the intersection kink point on the combined schedule comes from the group of individuals for whom, the combined schedule is just the upper bound of the AMT and regular income tax schedule – taxpayers who are unaffected by the long-term capital gains schedule. For this group, estimated elasticity is 0.20, as compared to 0.11 for taxpayers reporting some form of long-term capital gains.

Exploring differential bunching across time periods sheds light on taxpayers' reaction to the AMT across time. I conduct separate estimation of the elasticity of taxable income for time periods 1993-2002 and 2003-2011. 58 percent (42 percent) of the population represented by the analysis sample comes from the first (second) time period. The response to the intersection kink is greater for the latter time period as shown in Panels E and F of Figure 12. Estimated elasticity is 0.12 in time period 1993-2002, and 0.20 in time period 2003-2011.

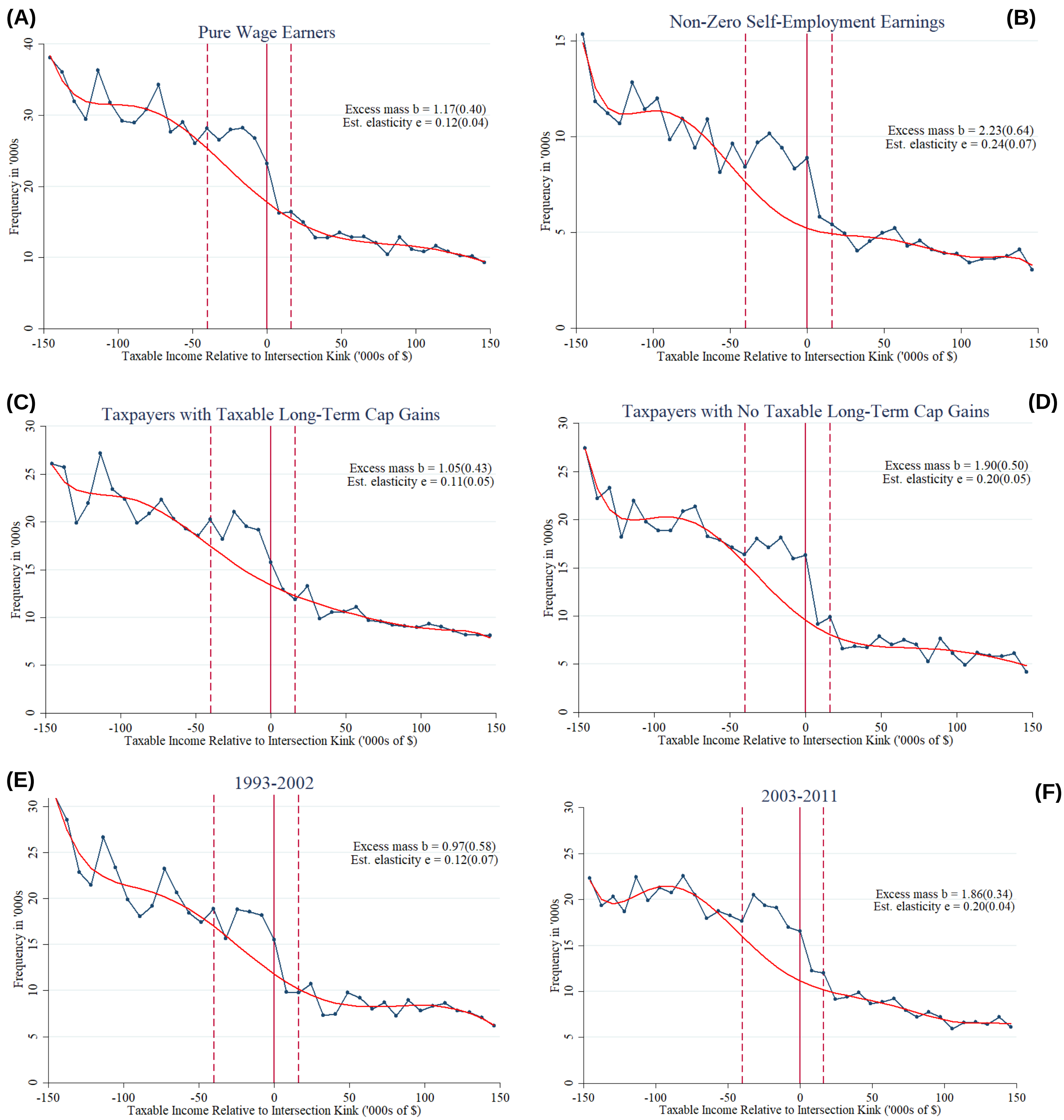


Figure 12: Distribution of Observed versus Counterfactual Effective Taxable Incomes (Disaggregated)

Notes: The figures here show the distribution of the observed and counterfactual densities for pure wage earnings (Panel A); taxpayers reporting any self-employment income (Panel B); taxpayers reporting long-term capital gains (Panel C); taxpayers reporting no long-term capital gains (Panel D); time period 1993-2002 (Panel E); and time period 2003-2011 (Panel F). The bold, vertical line (red) represents the centered intersection kink. Dashed, vertical lines (red) represent the lower and upper bounds of the bunching region defined as $-40,529 and $16,212. Observations are binned using a binwidth of $8,106. A seventh-order polynomial is used to construct the counterfactual distribution. Estimates for excess mass and the elasticity of taxable income with respect to the net-of-tax rate provided at the top-right for each figure. Bootstrapped SEs are provided in parentheses.

There are three institutional facets that can explain the difference in elasticity estimates across the two time periods. First, key features of the AMT schedule including the fixed deduction amount and marginal tax rates remained consistent from 1993 to 2000, with a slight increase in the fixed deduction amount for years 20001 and 2002. On the other hand, Congress increased the fixed deduction amount annually from 2003 to 2011, injecting uncertainty around the future AMT structure. Learning effects would predict that the inconsistency of the AMT schedule from 2003-2011 would lead to taxpayers optimizing behavior around the intersection kink with increased informational frictions. On the other hand, a stable policy environment allows taxpayers to gradually learn how best to optimize their economic and taxpaying behavior. This learning is also a function of the diffusion of information about a tax policy (Chetty, Friedman & Saez; 2013). Such diffusion plausibly slows down when policies change rapidly. Similarly, adjustment and search costs can also hamper optimization of real labor supply and tax strategies (Chetty, Friedman, Olsen, and Pistafferi, 2011). In terms of real outcomes, adjustment costs related to switching jobs (extensive margin) or altering hours worked (intensive margin) in response to rapidly changing marginal tax rates can be prohibitive (Gelber, Sacks & Jones; 2020). Optimizing tax strategies can also occur after a lag, such as shifting income across bases or time; and altering the structure of deductions, among others. If this is the case, then I would expect to find bunching behavior to be more pronounced in periods where the AMT and regular tax policies remained largely stable over time, and less so when policies changed more frequently. This would imply a lower bunching response to the left of the intersection kink.

Second, with the fixed deduction increasing substantially over the period 2003-2011, the intersection kink shifted to the right as compared to 1993-2011, thereby affecting individuals with higher earnings. Due to reasons discussed earlier, higher-income taxpayers have enhanced ability to change their behavior in response to changing marginal tax rates faced at the intersection kink. Therefore, it is possible that the bunching response would be higher in the second time period. Third, improvements in tax technology[18] used by taxpayers across time can reduce optimization frictions, leading to increased bunching in later years. Observed bunching responses shown in Panels E and F suggest that the aggregate effect of the second and third institutional features outweighs learning effects, leading to estimated elasticity in the second time period being higher than the first time period.

[18] "Tax technology" here refers to: digital tools which allow for flexible adjustment of inputs to assess the effect on tax liability of earning and reporting varying income levels across different streams; maturing of skills of tax accountants and advisors; and increases in information flow and access to such information on the internet.

Table 3 provides the average elasticity estimate for high earners, as well as estimates for different subpopulations. Worth noting are the higher elasticity estimates for all subpopulations in the second time period. In fact, the cleanest estimate for taxpayers unaffected by the additional complexity of the long-term capital gains schedule in later years is 0.28, much higher than the average elasticity estimate for high earners of 0.15. These results are reproduced and plotted in Figure 13.

Table 3: Summary of Elasticity Estimates for High Earners Around the Intersection Kink

Years	**MTR Change**	**All Filers**	**Self-employment Income**	**Wage earners only**	**Positive LTCG**	**Non-positive LTCG**
1993-2011	28% - 37.3%	0.15***	0.25***	0.12***	0.11**	0.20***
		(0.04)	(0.07)	(0.04)	(0.05)	(0.05)
1993-2002	28% - 39.5%	0.12*	0.25**	0.08	0.11	0.14
		(0.07)	(0.11)	(0.07)	(0.09)	(0.11)
2003 - 2011	28% - 35%	0.20***	0.26***	0.19***	0.13***	0.28***
		(0.04)	(0.07)	(0.04)	(0.04)	(0.06)

* Bootstrapped standard errors in parentheses
** Order of polynomial is 7
*** Binwidth = $8,106; Bunching Region: - $40,529 - $16,212
**** 1993-2002 covers two tax acts (OBRA and EGTRRA); 2003-2011 covers JGTRRA

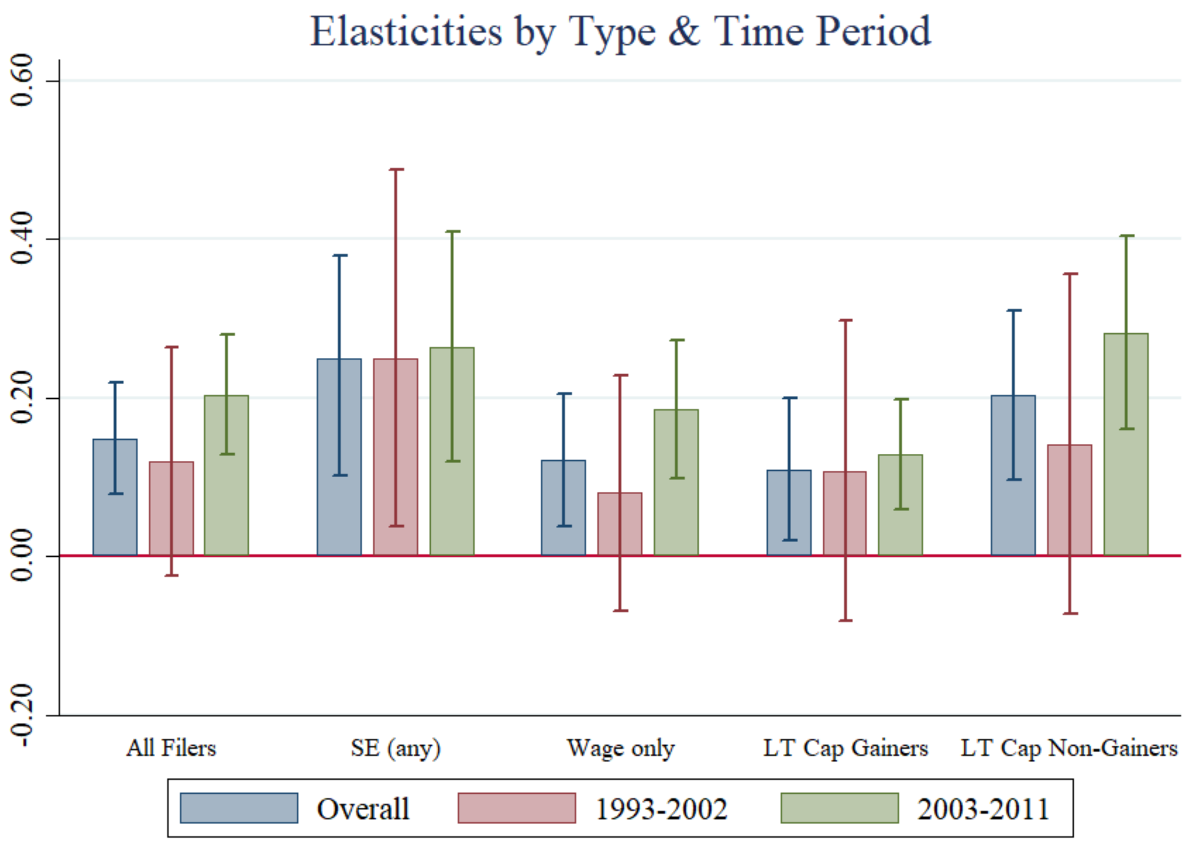


Figure 13: Elasticity Estimates for High Earners Disaggregated by Population Type

Notes: The figure illustrates the estimated ETI for high income taxpayers, disaggregated by type and time-period. The first bar in each case is related to the overall time-period; the second bar corresponds to 1993-2002; and the third bar is related to 2003-2011. The 95 percent confidence interval for each estimate is also provided.

VI. Sensitivity to Model Parameters and Robustness of Estimates

I test for the sensitivity of my average elasticity estimate to the choice of the following model parameters: binwidth and bandwidth. The choice of binwidth is a decision regarding the tradeoff between noise and bias. A smaller binwidth allows for capturing greater variation in the data that can be used to more precisely construct the observed and counterfactual taxable income densities, thereby reducing potential bias in the estimation of the elasticity parameter. On the other hand, the smaller the binwidth, the greater the variance. Increased variance adversely affects the ability to visually observe bunching behavior, and increases standard errors constructed for elasticity estimates. This is especially problematic if the total number of observations in the analysis window is not large.

Recall that I use IRS-provided sample of income tax returns data. For this sample, I focus on high income individuals. The fraction of high earners in the population is low. This is illustrated by

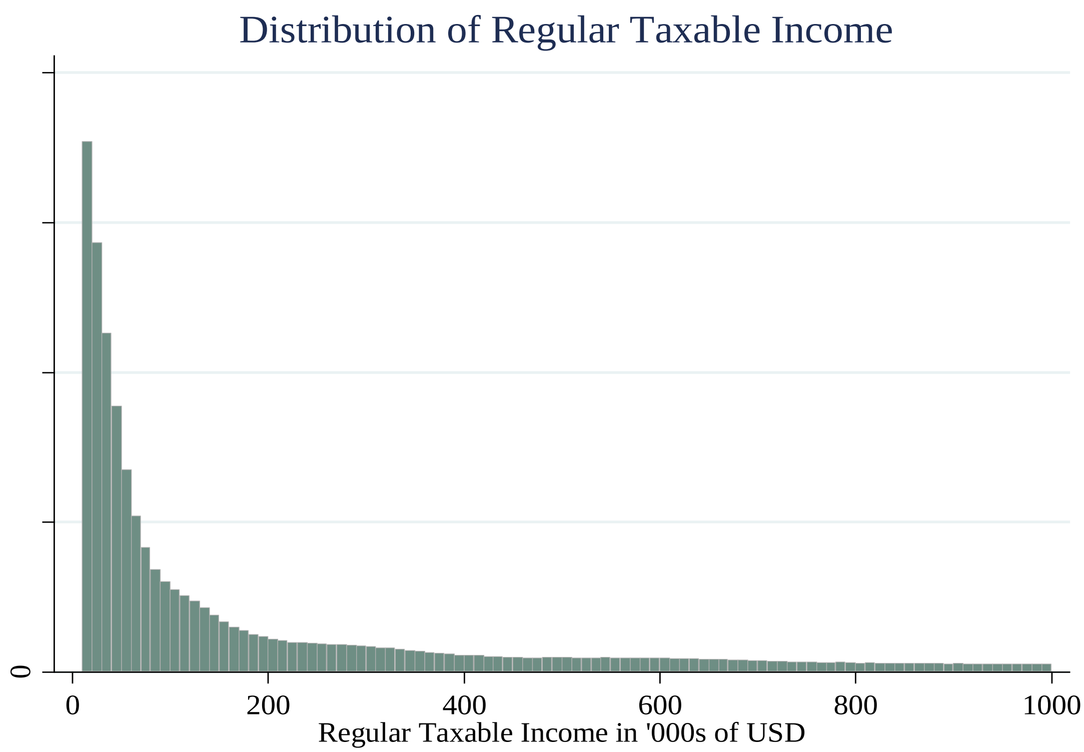


Figure 14: Distribution of Regular Taxable Income for the Time Period 1993-2011.

Notes: Binwidth is $10,000. The distribution is truncated below at $10,000 and above at $1 million.

Figure 14, which shows the distribution of regular taxable income, truncated at $10,000 and $1 million. It resembles a Pareto distribution, with a thin right-tail with few individuals at higher income levels, providing for fewer observations to work with in estimating elasticity. Further, my analysis window comprises the distribution of taxpayers with taxable incomes relative to the intersection kink point, leading to the analysis choosing a subset of these observations. Choosing a binwidth that is too small risks generating noisy estimates, with inference significantly affected by increased variance.

I use the Freedman-Diaconis rule to find the optimal binwidth of $8,106 for estimating the average elasticity of high earners of 0.15. However, I also assess the sensitivity of this estimate to varying binwidths. To do so, I choose binwidths between $500 and $12,000, reconstruct the counterfactual density and find estimates for the elasticity of taxable income corresponding to each binwidth. I plot these estimates and their 95 percent confidence intervals in Figure 15. It is reassuring to see that the average elasticity estimate for high earners remains stable over the range of binwidths considered.

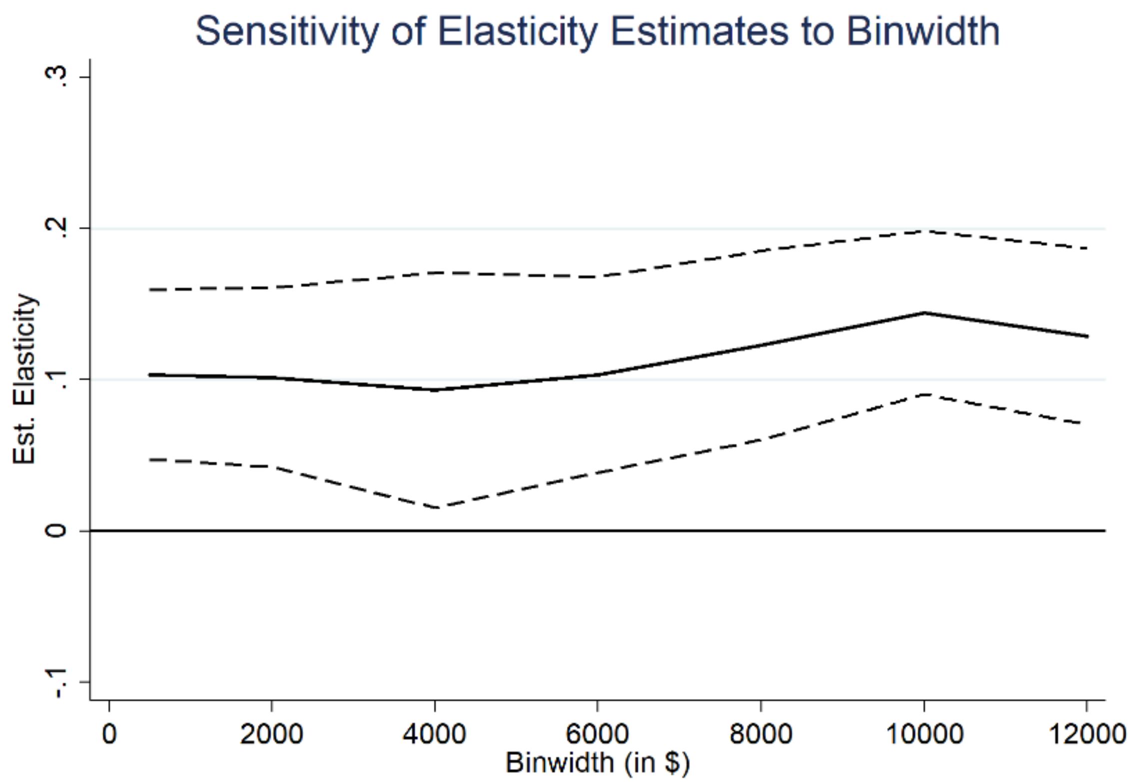


Figure 15: Sensitivity of Elasticity Estimates to Binwidth

Notes: This figure shows the sensitivity of the estimated ETI for high earners to the choice of binwidth. The solid curve represents the estimated ETI for each binwidth in $2,000 intervals. The dashed lines capture the confidence interval related to the estimated ETI across binwidth size.

The choice of bandwidth relates to the decision regarding the length of the bunching region considered. Unlike regression discontinuity designs where a smaller bunching region minimizes bias by ensuring that treatment and control groups on either side of a given cut-off are similar to each other, bunching estimators for measuring elasticity leverage the manipulation itself around the cut-off point. The cut-off point in the current context is the intersection kink point. If the bandwidth is too narrow, then manipulation revealed via bunching outside the assumed bunching window will not be captured, leading to elasticity being underestimated. On the other hand, a wide bandwidth might a) capture part of the distribution where no manipulation due to the kink point is taking place leading to underestimation of the elasticity parameter, or b) capture manipulation related to other parts of the tax schedule, such as other kink points leading to overestimation of the elasticity parameter.

Recent literature has placed bounds on the bunching region by exploiting visual bunching evidence. Besides utilizing the algorithm developed by Bosch et al. (2020), I also test the sensitivity of the average elasticity estimate for high earners to different bandwidths. Table 4 shows the results of this sensitivity check. A smaller bunching window corresponding to the first row of the table cuts into visually observed bunching, leaving out bunching behavior outside the assumed bunching window, potentially leading to elasticity being underestimated. An elasticity estimate of 0.12 for the smaller bandwidth, as opposed to the average elasticity estimate of 0.15 (middle row) indicates such underestimation. Increasing the bunching window, however, does not affect the average elasticity estimate. This makes sense, since if there is no strategic bunching in other parts of the distribution within the analysis window, or if other bunching cancels out, then increasing the bunching region would not impact the elasticity estimate.

Table 4: Sensitivity of Elasticity Estimates to Bandwidth

Years	MTR Change	Binwidth ($)	Bunching region ($)	All Filers
1993-2011	28% - 37.5%	8,106	-32,423, +16,211	0.12*** (0.03)
			-40,529, +16,211	0.15** (0.04)
			-48,634, +16,211	0.15*** (0.04)

* Bootstrapped standard errors in parentheses
** Order of polynomial is 7

Bertanha et al. (2018) devise a strategy to identify the elasticity of taxable income by assuming bounds on the magnitude of the slope of the heterogeneity probability density function (PDF), which is a Lipschitz continuity. By estimating the maximum slope magnitude of the PDF produces upper and lower bounds for all possible PDF values within the interval in which bunching occurs, in relation to the observed bunching mass. This leads to upper and lower bounds on the size of the bunching interval and therefore, upper and lower bounds on elasticity estimates. Estimating the bounds for the elasticity of taxable income with respect to the net-of tax rate, I find that the estimated lower bound is 0.12, while the estimated upper bound is 0.17. The average estimated elasticity of 0.15 computed using the traditional bunching estimator falls within these bounds.

VII. Efficiency Costs and the Optimal Top Marginal Tax Rate

The elasticity of taxable income parameter provides key insights into the responsiveness of taxpayers to changing marginal tax rates. By extension, it also serves as a critical parameter for estimating efficiency costs, and for conducting welfare analyses. In fact, assuming no externalities and market failures, and under negligible income effects, the elasticity parameter serves as a sufficient statistic for estimating efficiency costs of taxation (Feldstein, 1999). I return to the tenability of these assumptions at the end of this section.

Simplifying the model in terms of behavioral and mechanical costs of taxation, Saez, Slemrod and Giertz (2012) discuss how the literature has evolved to show that the marginal deadweight burden (MDB) or marginal excess cost of funds (MECF) is equal to $1 - dB/dR$, where dB is the extra amount of utility lost over and above additional tax revenue collected by a tax increase, while dR is overall change in tax revenue due to a tax increase. This translates to:

$$MECF(\tau, \varepsilon, \alpha) = \frac{1 - \tau}{1 - \tau - \varepsilon.\alpha.\tau}$$

Where τ is the prevailing top marginal tax rate, ε is the elasticity of taxable income, and α is the Pareto parameter. The right-tail of the income distribution can be shown to be Pareto distributed. The Pareto parameter α, estimates the thickness of the right tail. A thicker (thinner) right tail corresponds to a lower (higher) α. In relation to efficiency estimation, the thinner the tail and higher the Pareto parameter, the higher the loss of social utility on the margin relative to inframarginal revenue gains in the top tail due to tax rate increases, and the higher the efficiency cost. Efficiency costs increase in τ and ε, but decrease in α.

I use my estimates for the average elasticity for high earners, as well as the elasticity estimate for the sample unaffected by the additional complexity of the capital gains schedule in the second half of my analysis time-period to estimate efficiency costs using the above formula. I assume the α parameter to be equal to 1.5 based on the analysis of the US income distribution conducted by Piketty and Saez (2003). Similar to Saez et al. (2012), I assume average top state income tax rate to be 5.9 percent, Medicare payroll tax rate to be 2.9 percent, and average sales tax rate of 2.3 percent. The weighted average of the top marginal tax rate for my analysis period is 37.3 percent. Considering the deductibility of state income taxes from the federal income tax schedule, and the deductibility of the Medicare payroll tax from both state and federal income tax schedules, I estimate the average aggregate top marginal tax rate to be 44.8 percent.

$$MECF(0.448, 0.15, 1.5) = \frac{1 - 0.448}{1 - 0.448 - (0.15 * 1.5 * 0.448)} \approx 22\%$$

Estimation of efficiency costs using these parameters suggests that an additional dollar of income tax collected generates efficiency cost of 22 cents. Conversely, using the elasticity parameter for taxpayers with no long-term taxable capital gains in the more recent time period covering years 2003 to 2011, where the overall top marginal tax rate is 42.5 percent due to a lower federal top marginal tax rate of 35 percent, I estimate the efficiency cost to be 45 cents for the additional dollar collected in tax revenue.

$$MECF(0.425, 0.28, 1.5) = \frac{1 - 0.425}{1 - 0.425 - (0.28 * 1.5 * 0.425)} \approx 45\%$$

By combining efficiency costs with social welfare weights, the elasticity of taxable income with respect to the net-of-tax rate is also used for conducting welfare analyses. One needs an estimate of taxpayers' behavioral response measured via elasticity to be able to say something about the optimal top marginal tax rate. Building on the subliterature on optimal top marginal tax rates initiated by Mirrlees (1971), Diamond and Saez (2011) show that under the mechanical and behavioral effects on revenue of increasing taxation, and assuming away income effects with a quasi-linear utility function increasing in consumption but decreasing in effort, the optimal top marginal tax rate can be represented by:

$$\tau^* = \frac{1 - \bar{g}}{1 - \bar{g} + (\alpha * \varepsilon)}$$

Where $\bar{g}$ is the weighted average of the social marginal weights (g_i) for high earners. g_i can be thought of as the social marginal value of consuming an additional dollar relative to governmental revenue. Under a Rawlsian social welfare function where social marginal weights are concentrated at the bottom of the income distribution, $\bar{g} = 0$. Under a utilitarian social welfare function with concave utility functions, social marginal value of consumption for high earners decreases rapidly at the top of the income distribution, approaching zero. Therefore, the above formula is simplified to:

$$\tau^*(\alpha, \varepsilon) = \frac{1}{1 + (\alpha * \varepsilon)}$$

I plug in the Pareto parameter of 1.5 and my two main estimates of 0.15 and 0.28 for the elasticity of taxable income for high earners into the above and find corresponding optimal top marginal tax rates of 82 percent and 70 percent, respectively.

$$\tau^*(1.5, 0.15) = \frac{1}{1 + (1.5 * 0.15)} \approx 82\%$$

$$\tau^*(1.5, 0.28) = \frac{1}{1 + (1.5 * 0.28)} \approx 70\%$$

The range of estimated optimal top marginal tax rate coincides with emerging literature in the area (Diamond and Saez, 2011; Saez, Slemrod and Giertz, 2012; Piketty, Saez and Stantcheva, 2014). However, a caveat is in order. The formulae used to estimate efficiency costs and optimal top marginal tax rates here make relatively strong assumptions related to externalities, long-term responses to taxation, and the type of costs associated with taxation.

For example, increasing marginal tax rates along the income tax schedule can cause individuals to shift their income across tax bases in search of lower-taxed income streams, or across time. Such fiscal externalities (Saez et al., 2012) lead to some of the efficiency costs of income taxation to be recouped on other tax schedules or across time, leading to higher optimal top marginal tax rates. A similar argument can be made for classical externalities. Individuals increasing the use of tax avoidance via charitable giving or increased mortgage interest deduction amounts can generate externalities for other economic agents, reducing the efficiency cost of income taxation. Further, Chetty (2009) argues that if the costs of taxation are not purely real resource costs but instead include transfers to other agents as well – say, via tax penalties imposed for illegal tax avoidance or evasion, that are redistributed to other agents – then the elasticity of taxable income parameter is not sufficient for estimating efficiency costs. It is also possible that taxpayers retime income gain, leading to a wedge appearing between short-run and long-term elasticities, with income realized at a later period being taxed at a non-zero rate, affecting overall efficiency costs.

The study of these externalities is beyond the scope of this paper. Therefore, the estimation of efficiency costs and optimal top marginal tax rates should be used only as suggestive evidence. The core proposal made in the current analysis is the estimation of elasticities for high earners using the true tax incentive structure represented by the combined schedule, that comprises the interaction of the regular income tax and AMT schedules. Such measurement provides estimate of short-run elasticities which are different from zero, and allow for bounds to be placed on estimates of efficiency costs of taxing high earners.

VIII. Conclusion

The standard bunching estimation approach to measuring the elasticity of taxable income for high earners in the United States has been to construct the federal regular income tax schedule, overlay the distribution of taxable income across it, and then use taxpayer bunching responses around kink points to estimate the elasticity. This paper argues that the regular income tax schedule is not the correct schedule to use for estimating elasticity for high earners when bunching estimators are deployed. High-income taxpayers face an effective tax schedule that is the upper bound of the piecewise linear regular income tax and alternative minimum tax (AMT) schedules. This is the schedule that taxpayers respond to when optimizing behavior, and therefore, should form the underlying tax schedule used in bunching studies for higher earners. The use of the combined schedule for analysis of high earners' behavior resolves the inconsistency between previous empirical results (Saez, 2010; Mortenson and Whitten, 2016) and the standard static model in the literature relating kinks in the budget set and bunching response of taxpayers that predicts higher responses for high income taxpayers.

I characterize this combined schedule and reveal its properties, allowing me to capture larger bunching responses, as well as to mitigate endogeneity concerns plaguing earlier bunching studies. The combined schedule contains its own kink points which do not necessarily align with kinks on the regular income tax schedule when the latter is considered in isolation. In fact, the intersection kink on the combined schedule – the point where the regular income tax and AMT schedules intersect – provides a novel device for measuring taxpayer response. The intersection kink presents a larger change in the marginal tax rate, plausibly generating greater behavioral response as compared to top kinks on the regular income tax schedule. Further, the location of the intersection kink varies for each taxpayer, as compared to kinks on the regular income tax schedule that are fixed in taxable income. This variation allows for elasticity estimation to be more robust to bias plaguing earlier studies, by disentangling variation in taxable income from variation in underlying preferences.

Using publicly available IRS taxpayer microdata from 1993-2011, I find that the average elasticity estimate for high earners in the United States is 0.15, as compared to earlier estimates in the literature that were close to zero. For the sample unaffected by the complexity of the capital gains schedule, the estimated elasticity is 0.20 – rising to 0.28 for the time period 2003-2011 when annual changes in the tax code shifted intersection kinks for taxpayers to higher parts of the income distribution, on average. Self-employed individuals respond twice as much as wage earners, with an elasticity of 0.24. However, wage earners also reveal non-trivial responsiveness to tax rates, with an estimated elasticity of taxable income of 0.12, potentially shedding light on the enhanced ability of taxpayers at the top of the income distribution to alter work hours, or to convert monetary compensation to fringe benefits. Back of the envelope calculations reveal efficiency costs ranging between 22 cents to 45 cents per dollar of additional tax revenue collected and estimated optimal top marginal tax rates between 70 percent and 82 percent.

The current analysis creates a range of avenues for future work. In the context of the United States, further analyses should look at the interaction of the regular and alternative minimum tax schedules together with the capital gains schedule. While this additional complexity increases the dimensionality of the problem, capturing the shape of the combined schedule that also incorporates the capital gains schedule is important for constructing the correct underlying schedule. A related issue is the differential treatment of tax credits, specifically the foreign tax credit on the regular income tax and AMT schedules. Studying the impact of this tax credit on the location of the intersection kink and ensuing bunching behavior would provide greater insights into taxpayer behavior. Future work should also consider dynamic responses of taxpayers by using tax panels available at the IRS, to a) unpack the mechanisms underlying individuals' responses to the combined schedule across years, and b) to shed light on short-term versus long-term responses to the combined schedule.

Beyond the United States, this paper expounds the importance of considering details of the tax code that give rise to combined schedules with characteristics that are different from the primary tax schedule being considered. Such under-the-hood work is necessary for identifying the true incentive structure faced by taxpayers, when estimating taxpayer responsiveness to such incentives.

Bibliography

Angrist, Joshua D., and Alan B. Krueger, 1992. “The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables With Moments From Two Samples”, Journal of the American Statistical Association, 87,328-336.

Best, Michael C., Brockmeyer, Anne, Kleven, Henrick J., and Mazhar Waseem, 2015. “Production vs Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan”, Journal of Political Economy 123(6), 1311-1355

Bruce, Donald, and Xiaowen Liu, 2014. “Tax Evasion and Self-Employment in the US: A Look at the Alternative Minimum Tax”, Internal Revenue Service, Statistics of Income Division Publication

Burman, Leonard E., Engler, Alex, Khitatrakun, Surachai, Nunns, James R., and Sarah Armstrong, 2017. “A Synthetic Income Tax Return Data File: Tentative Work Plan and Discussion Draft”, Tax Policy Center, Urban Institute and The Brookings Institution

Chetty, Raj, Friedman, John, Olsen, Tore, and Luigi Pistaferri, 2011. “Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records”, Quarterly Journal of Economics, 126, 749–804

Dahl, Molly, Schwabish, Jonathan, DeLeire, Thomas, and Timothy Smeeding, 2012. “e Earned Income Tax Credit and Expected Social Security Retirement Benefits Among Low-Income Women”, Congressional Budget Office, Working Paper 2012-06

Emmanuel Saez, 2010. “Do Taxpayers Bunch at Kink Points?”, American Economic Journal: Economic Policy 2 (August 2010): 180–212

Feenberg, Daniel R., and James M. Poterba, 2004. “The Alternative Minimum Tax and Effective Marginal Tax Rates.” National Tax Journal 57(2), 407–427

Fleenor, Patrick, and Andrew Chamberlain, 2005. “Backgrounder on the Individual Alternative Minimum Tax (AMT)”, Tax Foundation

Gruber, Jon, and Emmanuel Saez, 2002. “The Elasticity of Taxable Income: Evidence and Implications”, Journal of Public Economics, 84(1), 1-32.

Hamermesh, Daniel S., Frazis, Harley, and Jay Stewart, 2005. “Data Watch: The American Time Use Survey”, Journal of Economic Perspectives, Volume 19, Number 1, Winter 2015, Pages 221–232

Internal Revenue Service, Statistics of Income (SOI) Division, “Design Changes to the SOI Public Use File (PUF)” (https://www.irs.gov/pub/irs-soi/14rfpufredesignrecommen.pdf)

Internal Revenue Service, Statistics of Income (SOI) Division, 2010. “SOI Products and Services” (https://www.irs.gov/pub/irs-soi/12otproductswinbul.pdf)

Inoue, Atsushi, and Gary Solon, 2010. “Two-Sample Instrumental Variables Estimators”, The Review of Economics and Statistics 2010 92:3, 557-561

Kleven, Henrik J., Knudsen, Martin B., Kreiner, Claus T., Pedersen, Soren, and Emmanuel Saez, 2011. “Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Demark”, Econometrica, Vol. 79, No. 3, 651–692

Kleven, Henrik J., and Mazhar Waseem, 2013. “Using Notches to Uncover Optimization Frictions and Structural Elasticities: Evidence from Pakistan”, Quarterly Journal of Economics, 128, 669–723, 2013

Leiserson, Greg, and Jeffrey Rohaly, 2008. “The Individual Alternative Minimum Tax: Historical Data and Projections”, Tax Policy Center, Urban Institute and The Brookings Institution

Mazhar Waseem, 2017. “Using Movement of Exemption Cutoff to Estimate Tax Evasion: Evidence from Pakistan”, Review of Economics and Statistics (Revise & Resubmit)

Mike Weber (Presentation), 2009. “Individual Income Tax Return Panel Data at SOI: What SOI Has and How You Can Use It”, IRS SOI (https://www.bus.umich.edu/otpr/papers/Weber.pdf)

National Bureau of Economic Research, 2017 “US Individual Income Tax Public Use Sample Documentation” (http://users.nber.org/~taxsim/gdb/)

Pissarides, Cristopher, and Guglielmo Weber, 1989. “An Expenditure-Based Estimate of Britain’s Black Economy”, Journal of Public Economics 39(1), 17–32.

Saez, Emmanuel, Slemrod, Joel, and Seth H. Giertz, 2012. “The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review”, Journal of Economic Literature, 50:1, 3–50

U.S. Census Bureau, 2017. “American Time Use Survey User’s Guide”, (https://www.bls.gov/tus/atususersguide.pdf)

Appendix

Table A: Previous Estimates of the Elasticity of Taxable Income for High Earners

Author(s)	**Data (Years)**	**Tax Change**	**Income Definition**	**Methodology**	**Elasticity Estimate**
Panel A: Regression-based, Difference-in-differences, Instrumental Variable					
Lindsey (1986)	1979-1984	ERTA 81	Taxable Income	DiD	1.7
Feldstein (1995)	1985, 1988	TRA 1986	Taxable Income	DiD	1.25 - 2.14
Sammartino and Weiner (1997)	1989-1995	OBRA 1990 and OBRA 1993	AGI	DiD	Close to zero permanent AGI response
Carroll (1998)	1989-1995	OBRA 1991 and OBRA 1993	Taxable income	Regression-based	0.4
Moffitt and Wilhem (1998)	1983, 1989	TRA 1986	AGI	DiD; FE; IV	0 to 2 (depending on instruments)
Goolsbee (2000)	1991-1995	OBRA 1993	Taxable Income	Fixed Effects and First Differencing	Short-run > 1; Long-run: 0 - 0.4
Gruber and Saez (2002)	1979-1990	Changes across pairs of years	Taxable income	IV Approach	0.57
Saez (2003)	1979-1990	Bracket creep 1979-1981	Taxable Income	Instrumental Variable Approach	0.3 (not statistically significant)
Saez (2017)	2011-2015	2013 top bracket tax rate increase	Taxable income	Comparison of income shares	Short-run: 1.16; medium-run: 0.26
Panel B: Bunching					
Saez (2010)	1988-2004	Income Tax Kinks	Taxable income	Bunching	0.006 to 0.031
Mortensen and Whitten (2016)	1996-2016	Income Tax Kinks	Taxable Income	Bunching	0 (no response)
Chetty et al. (2011) ***	1994-2011	Income Tax Kinks	Taxable Income	Bunching	0.02
Kleven et al. (2011) ***	2007-2008	Income Tax Kinks, Audit Probability	Taxable income	Bunching, Experimental	Self-employed: 0.16; Stock income: 2.24
Kleven et al. (2014) ***	1991-2010	Danish preferential foreigner tax scheme	Annualized Earnings	Bunching	0.3

**** Studies indicated with asterisks are based on income tax data from Denmark*

Annex A: Legislative History

Policy Instrument (Year)	Description
Tax Reform Act of 1969 (P.L. 91-172)	Introduced the “add-on” minimum income tax of 10% in excess of an exemption of $30,000.
Excise, Estate, and Gift Tax Adjustment Act of 1970 (P.L. 91-614)	Allowed deduction of the “unused regular tax carryover” from the base for the minimum tax.
Revenue Act of 1971 (P.L. 92-178)	Imposed minor provisions regarding foreign income.
Tax Reform Act of 1976 (P.L. 94-455)	Raised rate of minimum income tax to 15% and lowered exemption to $10,000 or half of regular taxes.
Tax Reduction and Simplification Act of 1977 (P.L. 95-30)	Reduced minimum tax preference for intangible costs of drilling oil and gas wells.
Revenue Act of 1978 (P.L. 95-600)	Introduced AMT alongside minimum income tax and moved certain itemized deductions and capital gains to AMT. AMT had graduated rates of 10%, 20%, and 25%, and an exemption of $20,000.
Economic Recovery Tax Act of 1981 (P.L. 97-34)	Lowered AMT rates to correspond with reductions in rates of regular income tax.
Tax Equity and Fiscal Responsibility Act of 1982 (P.L. 97-248)	Repealed “add-on” minimum tax. Made AMT rate a flat 20% of AMT income after exemptions of $30,000 for individuals and $40,000 for joint returns.
Deficit Reduction Act of 1984 (P.L. 98-369)	Made minor changes concerning investment tax credit, intangible drilling costs, and other items.
Tax Reform Act of 1986 (P.L. 99-514)	Raised AMT rate to 21%. Made high-income taxpayers subject to phase-out of exemptions. Increased number of tax preferences. Allowed an income tax credit for prior year AMT liability.
Revenue Act of 1987 (P.L. 100-203)	Made technical corrections related to Tax Reform Act of 1986.
Technical and Miscellaneous Revenue Act of 1988 (P.L. 100-647)	Made technical corrections related to Tax Reform Act of 1986.
Omnibus Budget Reconciliation Act of 1989 (P.L. 101-239)	Made further technical amendments.
Omnibus Budget Reconciliation Act of 1990 (P.L. 101-508)	Raised AMT rate to 24%.
Energy Policy Act of 1992 (P.L. 102-486)	Changes regarding intangible costs of drilling oil and gas wells.
Omnibus Reconciliation Act of 1993 (P.L. 103-66)	Introduced graduated AMT rates of 26% and 28%. Increased exemption to $33,750 for individuals and $45,000 for joint returns. Changed rules about gains on stock of small businesses.
Taxpayer Relief Act of 1997 (P.L. 105-34)	Changes regarding depreciation and farmers’ installment sales.
Tax Technical Corrections Act of 1998 (P.L. 105-206)	Adjusted AMT for new capital gains rates.

Tax Relief Extension Act of 1999 (P.L. 106-170)	Changed rules about nonrefundable credits.
EGTRRA (2001)	Tax Cuts and No change in AMT
2006	Introduction of calculator
American Taxpayer Relief Act of 2012	Indexes to inflation the income thresholds for being subject to the tax
2001-2012	Changes in Exemption Rates

Annex B: Exemption Rates Across Time

Years	Individual tax rate	Married filing jointly ($)	Single or head of household ($)
1986–1990	21%	40,000	30,000
1991–1992	24%		
1993–2000	26% / 28%	45,000	33,750
2001–2002		49,000	35,750
2003–2005		58,000	40,250
2006		62,550	42,500
2007		66,250	44,350
2008		69,950	46,200
2009		70,950	46,700
2010		72,450	47,450
2011		74,450	48,450
2012		78,750	50,600
2013		80,800	51,900
2014		82,100	52,800
2015		83,400	53,600
2016		83,800	53,900
2017		84,500	54,300
2018		86,200	55,400

Annex C: Exemption Rates and Phase-Out in the Early 2000s

Status	Single	Married filing jointly	Married filing separately	Trust	Corporation
Tax Rate: Low	26%*	26%*	26%*	26%*	20%*
Tax Rate: High	28%*	28%*	28%*	28%*	20%*
High Rate Starts (2012 and earlier)	$175,000	$175,000	$87,500	$175,000	n/a
High Rate Starts (2013)	$179,500	$179,500	$89,750	$179,500	n/a
Exemption in 2009	$46,700	$70,950	$35,475	$22,500	$40,000
Exemption in 2010	$47,450	$72,450	$36,225	$22,500	$40,000
Exemption in 2011	$48,450	$74,450	$37,225	$22,500	$40,000
Exemption in 2012	$50,600	$78,750	$39,375	$22,500	$40,000
Exemption in 2013	$51,900	$80,800	$40,400	$23,100	$40,000
Exemption phase-out starts at (2012 and earlier)	$112,500	$150,000	$75,000	$75,000	$150,000
Exemption phase-out starts at (2013)	$115,400	$153,900	$76,950	$76,950	$150,000
No more exemption in 2009 at	$299,300	$433,800	$216,900	$165,000	$310,000
No more exemption in 2010 at	$302,300	$439,800	$219,900	$165,000	$310,000
No more exemption in 2011 at	$306,300	$447,800	$223,900	$165,000	$310,000
No more exemption in 2012 at	$314,900	$465,000	$232,500	$165,000	$310,000
No more exemption in 2013 at	$323,000	$477,100	$238,550	$165,000	$310,000
Long-term capital gains rate	15%	15%	15%	25%	20%

** For income within the exemption phase-out, marginal tax rates are effectively multiplied by 1.25, which changes 20% to 25%, changes 26% to 32.5%, and changes 28% to 35%.*